

How to make a stronger NPT

Worries about nuclear weapons and their manufacture, dramatized by anxiety about Iraq, can be met only by strengthening the Non-Proliferation Treaty, due to expire in 1995. There is not much time left.

WHEN will the next nuclear war break out? And how will it begin? It is strange that these questions have become more immediate so soon after the ending of the Cold War, and in the wake of bilateral treaties between the Soviet Union (as it used to be) and the United States, but that assessment seems fair. The revelations by safeguards specialists from the International Atomic Energy Agency (IAEA), as well as at the Board of Governors' meeting in the past weeks, show that Iraq had gone much further towards the development of nuclear weapons than even Israeli intelligence had been saying; by all accounts, a factory for producing isotope-separation centrifuges had begun operation, while now there is a hunt for a supposed plutonium-producing reactor.

If Iraq is now discovered to be so far advanced on this sinister path, what can be said about Pakistan, whose ostensibly nuclear programme goes back to the 1960s, about North Korea (which has announced almost publicly that it has a reactor in operation without safeguards), South Africa and, for that matter, Israel? And then there is the newly generated legitimate concern about the future of the Soviet stockpile of nuclear weapons; will there remain a central defence organization to take charge of them, or will they finish up in the hands of whichever centrifugal republics make successful bids for them?

Ambitions

The case of Iraq is exceptional in one sense only. Before the invasion of Kuwait last year, Iraq was blessed with a stream of state-controlled oil revenues that made it the envy of many of its neighbours. The government did not hide its military ambitions, fanned by the long war against Iran, and could easily have diverted some thousands of millions of dollars towards a nuclear programme. The credit of \$3,000 million won illicitly from the Atlanta branch of the *Banco Romano del Lavoro* in 1987 would not have hindered the development, while the confessed objective of the invasion of Kuwait was also monetary. It is difficult to believe that any other of the candidate nuclear powers has been able to throw as much money at the problem.

But there are countervailing arguments. Time and patience are good substitutes for mere money, while Iraq's example has shown that skill in nuclear manipulation is readily learned or acquired. (The days are long past when established nuclear powers could safely patronize even

the most unprepossessing states by supposing them incapable of emulation.) That is why it is now prudent to suppose that all of the countries which have not so far signed the Nuclear Non-Proliferation Treaty (NPT) are intrinsically nuclear powers — as are even some of the signatories (among whom was Iraq).

Dangers

How dangerous would that make the world in which we live? The comprehensiveness of Iraq's nuclear programme, with no fewer than three independent routes to fissile material, coupled with its advertized interest in long-range missiles and even its supergun, point only to one goal: a nuclear attack on Israel. But how could a smallish country, even with an army comprising some 20 per cent of the adult population, hope to get away with such a barefaced assault on international civility and yet escape the threat of retaliation from Israel's protector in the Middle East, the United States?

There are trains of thought that might make the gamble seem worthwhile. Prudently, a would-be attacker might provoke (or arrange to provoke) Israel into a conventional war with its neighbours and hope to win international sympathy before making a move itself. Or it might hope to launch such a devastating first strike that the prospect of retaliation could be presented to the outside world as a wanton piece of bullying. Both calculations are, of course, uncertain, but the courses of action they might have impelled are hardly the less destructive on that account. And while there are powers that must be regarded, for practical purposes, as nuclear powers, similar dangers will continue to recur.

That is why the NPT, which expires unless renewed in 1995, should now be renegotiated urgently. Recruiting new signatories of the treaty in its present form, while desirable, is less important than arranging for a treaty with greater force. And the task should not be impossible.

The major nuclear powers have, indeed, made substantial progress towards the control of strategic arms (which they had promised in the NPT) and could, if events in the Soviet Union follow a reasonable course, be persuaded quickly to make more. The asymmetry between nuclear and non-nuclear powers was one of the obstacles to the negotiation of the NPT in the late 1960s. And then, if greenhouse calculations are correct, the years ahead are likely to see a revival of interest in civil nuclear power and



an emerging need in the poorer countries with technical assistance in this field — a development that could breathe new life into the dead-letter promises by the nuclear powers in 1970. The framework of the NPT could be kept alive.

How should it be strengthened? There is some relatively modest tightening of the NPT that needs to be accomplished. Despite the existing agreement between the nuclear suppliers that they will not assist others to make nuclear weapons, the case of Iraq shows that the agreement leaks. (Finding out which companies and people helped with the developments of the past few years is rightly part of IAEA's objective; their identities should be made public.) There is a strong case for greater formality, and for some kind of public register of supplies of nuclear equipment along the lines that the prime minister of Japan has been advocating for arms sales in general. The case for sanctions against illicit suppliers and their customers deserves a hearing.

The more desirable end-point is more drastic: a general understanding within the United Nations that the novel development of nuclear weapons is an offence against the international community, and must be stopped (as will, eventually, be the development in Iraq). That would require a decision by the UN Security Council that non-signatories of the amended NPT must involuntarily submit their nuclear installations to international safeguards. If the Security Council decided as recently as a year ago that the invasion of Kuwait must be undone, why should not it take this further step in this direction?

There are snags. For one thing, it would be unseemly that the established nuclear powers should ask for such a regime without making like sacrifices themselves, but general inspection of all nuclear facilities is not now an insupportable prospect. Indeed, might not this be the time to revive the old idea of a cut-off of military production (France's long-standing condition for signing the NPT, itself now softened)? Obdurately silent China is one stumbling-block. So, too, is Israel, which would be automatically denied US assistance of any kind if it were shown to be a nuclear power, but which will dig in its heels unless this year's peace conference on the Middle East yields sweeter fruit than is now in prospect. The present danger is that these obstacles, seeming insurmountable, will paralyze thought about a stronger NPT. Rather, they should be a spur to its reform. □

Straight-line city

The case for a linear city to the east of London is stronger than the short-term arguments suggest.

MR Michael Heseltine, who might well now have been the British Prime Minister had not Mrs Margaret Thatcher withdrawn from the second round of polling to elect the leader of the Conservative Party, has — in his lesser role as Secretary of State for the Environment — a splendid idea: he wants to create a linear city stretching eastwards

from the City of London along the Thames for perhaps 50 miles to the coast of the North Sea. There appear to be two short-term incentives — to breathe new life into development projects mounted in the past few years immediately to the east of London and now languishing in the worldwide property boom, and to divert development in Britain away from southeast England (full of articulate and influential protesters at machines shaped like bulldozers) towards the ugly and largely unloved north bank of the estuarine Thames. But the notion should not be crabbed on that account. The advantages of linear cities are unrecognized, let alone realized.

The reasons why traditional inner cities have become slums are well-recorded. People who live in cities seek occasionally to escape from them; the better off they are, the more readily they move to suburbs. But in a linear city, downtowners can escape laterally if they wish. (Heseltine's city would offer escape towards either a marine environment or dry land.) That is the crucial consideration. The price that must be paid is that of effective high-speed transport along the backbone of the city. City-dwellers could not call themselves that if it were to take more than, say, an hour to get from any one place to any other, which argues for a backbone of several high-speed railway tracks. Indeed, if Heseltine were to persuade his government to install the tracks, he would find his linear city springing into being without further fuss. □

Telephone silence

Academics in the United States, where the telephone was invented, are over-secretive about their telephone numbers.

THE telephone is a remarkable invention which has helped to change the world. As now developed, it enables people to speak to others at the far ends of the Earth, usually when they need to do so. That is why a person's telephone number is often more valuable than his or her street address. And that is why the failure of many scientists to include their telephone (and fax) numbers on their letter-heads is a constant source of frustration and perplexity, especially to a journal such as *Nature*.

The worst offenders are researchers at institutions in the United States, particularly at the various campuses of the University of California. Why this reticence? Is it that, with the virtual disappearance of the personal assistant, researchers are shy of receiving telephone calls unannounced? (The personalized answering machine is the defence against that.) Or shy of revealing exactly where they are at any time? (One Californian researcher who claims to divide his time between three telephone numbers has answering machines that refer callers to the alternative numbers, often fruitlessly.) Whatever the explanation, it is indeed strange that one of the most technologically advanced societies in the world should not yet have come to terms with the invention of the telephone. And is it not doubly ironic that the telephone should have been invented in the United States? □

Department of Energy faces hard choices

- Group advises postponing, cancelling projects
- Main Injector, fusion experiment are victims

Washington

TO DEAL with a declining budget for basic research, the Department of Energy (DOE) should postpone or cancel several major projects that physicists had hoped to begin in the next couple of years, including the 'burning plasma experiment' in fusion research and two planned high-energy physics facilities, the 'Main Injector' at Fermilab and a 'B Factory' at the Stanford Linear Accelerator Center (SLAC). That was the consensus of a group of senior researchers assembled last week to advise William Happer, DOE's newly appointed director of energy research.

Happer, who was confirmed by the Senate on 2 August, was quickly faced with a near-crisis. According to recent projections, DOE budgets for research through 1996 are expected to decline in constant dollars — exclusive of the Superconducting Super Collider (SSC), which is a separate item in the federal budget. At the same time, mounting environmental and safety costs are steadily forcing up the price of research. Something had to give.

And it had to give quickly. Happer's Office of Energy Research must offer a fiscal 1993 budget to DOE within the next few weeks that describes its priorities for allocating research funds. So Happer called for help. He assembled a task force of senior researchers and research administrators to advise him on what should go and what should stay. He asked them not only to make suggestions on priorities within a field — how the high-energy physics budget should be allocated, for example — but also for advice on how to divvy up the research budget across fields. Happer indicated that he is likely to follow the advice of the committee closely.

The committee reached a consensus on funding priorities in each of the major areas of research funded by the DOE:

■ **High-energy physics.** DOE should continue its commitment to building its flagship project, the Superconducting Super Collider. The department should not proceed at this time with the Main Injector at Fermilab, a proton synchrotron that is designed to replace the lab's Main Ring, although there is already money to start its construction in the fiscal 1992 budget; the DOE should ask the High-Energy Physics Advisory Panel whether it still wants to proceed with the Main Injector, given the changed spending constraints. No decision should be made on the proposed B Factory, an electron-positron collider designed to study CP violation. The SSC

doesn't completely escape cuts, either. The committee recommended that funds should not be set aside specifically for SSC Laboratory scientists to do research at other facilities until the SSC is finished, with the goal of developing expertise at the new lab; instead, the SSC Laboratory scientists should compete with other high-energy physicists on an even footing.

■ **Fusion research:** The burning plasma experiment, a \$1,400 million, 10-year project that was planned as the next step in US fusion research, should be cancelled. The committee did agree, however, that because fusion research has been hit hard by a declining budget over the past decade, the DOE should try to increase the fusion budget. The department should consider collaborations on foreign fusion projects as an intermediate measure until the International Experimental Test Reactor (ITER) comes on line sometime in the first decade of the next century.

■ **Materials research:** The design effort for the Advanced Neutron Source should be continued, and the materials community should be asked whether the department should fund that project at the cost of other programmes. The Advanced Photon Source, now under construction at Argonne National Laboratory, should be continued as planned.

■ **Hadron physics:** The Relativistic Heavy Ion Collider is the main priority for this area of physics and is now being built at Brookhaven National Laboratory. Its construction will squeeze other parts of the hadron physics budget, however, and the committee came to no solid conclusion as to how to handle this.

In general, the task force recommended that advisory groups from the different areas of physics meet and tell DOE what their own priorities are, given that they have much less money to spend than originally thought. The committee made a few generalizations about the allocation of funds across fields, but for the most part seemed to agree with Charles Baltay of Yale University, who said, "We should not fiddle lightly with the balance between fields." Nonetheless, with increasing pressure on its budget, the DOE is likely to make some of those hard choices in the next year or two, and decide to put more money into one area of physics at the expense of another. The committee members agreed, however, that it will take more than a two-day meeting to make those picks.

Robert Pool

New findings cast doubt on UK vaccine trials

PROMISING trials in Britain of a vaccine to protect against AIDS-like symptoms in monkeys have been derailed by the news that the encouraging results to date were more likely due to experimental artefacts than to signs of general immunity. (See also page 297.)

Until the announcement at a Medical Research Council (MRC) AIDS workshop last week, trials at the National Institute for Biological Standards and Control had suggested that a vaccine could protect macaques against the simian immunodeficiency virus (SIV). Macaques had been inoculated with 'killed', or inactivated SIV, allowed to develop antibodies, and then 'challenged' with live SIV. Many of them appeared to produce antibodies to the live virus, which prevented the virus from entering the body's cells and protected the monkeys from the simian equivalent of AIDS.

But the new results suggest that the monkeys were not reacting only to the SIV, as had been thought, but also to components of the human cell line in which the virus was grown. Both the virus that is inactivated before it is injected in the monkeys and the live challenge virus were grown in the same human T-cell lines. In the MRC trial, four macaques were given a vaccine based on inactivated human T cells that had been infected with SIV. Another four macaques were inoculated with inactivated T cells that had not been infected.

The researchers challenged both sets of macaques, and found that three of four that had been given the 'real' vaccine were protected. When they examined the other group, however, they were shocked to discover that half of them produced antibodies to the virus as well, despite the fact that the vaccine they had been inoculated with had never seen SIV, either alive or dead.

In a correspondence to be published in *Nature* next week, E. J. Stott and his MRC colleagues say that the findings do not rule out the possibility that the monkeys are also developing antibodies to the virus itself. But the mechanism for such protection is still far from clear. The MRC team plans to check its results by using a virus of a sort that can be propagated in simian peripheral blood lymphocyte (PBL) cells. That way, only the virus would be registered as foreign to the monkey's immune system, and any antibody reaction can be safely assumed to be SIV-related. If those trials trigger no immune responses, then the results to date must be considered "extremely suspect", says MRC project leader Peter Greenaway. "Then we've got to go back and consider the entire spectrum of approaches [to an AIDS vaccine] again."

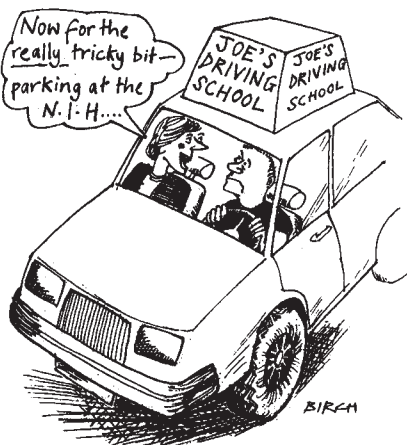
Christopher Anderson

Culture clash inside the walls

Bethesda, Maryland

FROM the outside, the US National Institutes of Health (NIH) look like a troubled agency. Scientific misconduct, conflict of interest, funding shortages, White House meddling on fetal tissue research and a cancelled sex survey have all made headlines this year. Inside, things look different, but not necessarily better.

Last week, NIH researchers got their chance to sound off, when director Bernadine Healy held the agency's first 'town meeting' in hopes of avoiding a



growing crisis of morale. The agency's researchers turned out to be no happier about NIH than the average congressional investigator, but that was where the similarity ended. Paperwork, parking, and laboratory space are the things that keep intramural researchers up at night, not data fabrication and politics.

Indeed, congressional concern about government contracting abuses and con-

flict of interest have only made life more difficult at NIH, researchers complained. Paperwork and bureaucracy often consume a third of an NIH researcher's time. One example: NIH scientists are supposed to buy equipment from small companies whenever possible. A year ago, it was enough to note that one had checked with such a small company before going to a larger competitor. But last month a memorandum informed staff of a new rule: regardless of what their catalogues may indicate, at least two small companies must be telephoned and stock clerks questioned (and their names taken down) before it would be acceptable to go to a large company for a hard-to-find item.

"We are treated like naughty children who are wont to tell lies," complained yeast researcher Enrico Cabib. "In 20 years at NIH, I think I have earned the trust and respect of the scientific community. It is sad that I have not earned the trust of the people where I work."

Other researchers suggest that, rather than acquiescing while Congress passes new laws to regulate NIH, perhaps the agency could hire some good lawyers to get around those laws it is already saddled with. And nearly everybody mentioned the space problem, whether it be no place to put their car or no room but the hallway for their laboratory's autoclave. One researcher told of a colleague who worked on a plank over a sink in the crowded Bethesda campus before he was gratefully transferred to 'Siberia' — the NIH satellite facility in Frederick, Maryland. At least there he had his own workspace, even if he did have to drive an hour to get to it.

Dingell to Healy: lighten up

BERNADINE Healy can rest easy for the moment. Whatever else may come of her running battle with the House Investigations and Oversight subcommittee under John Dingell (Democrat, Michigan), the NIH director will not have to go jail over her parking space.

Healy herself raised that spectre at the NIH 'town meeting' last week, after staff researchers challenged her over parking space shortages. Although she promised some 600 new spaces over the next six months, Healy pointed out that there are some things worse than having no space at all — a congressional investigation, for example. Dingell's staff descended upon NIH last month to investigate her handling of misconduct and other matters, she revealed, "and one of the things they looked into was my parking space."

That seems a bit extreme, even for the detail-obsessed Dingell. A call to the

subcommittee shed some light on the subject: during their visit, Dingell investigators Peter Stockton and Bruce Chafin mistakenly parked in Healy's spot, prompting one of her staff to run out and shoo them away. Why, the investigators teased the aide, does Healy need a parking space, anyway? After all, she lives in a government-supplied house on the NIH campus, a few hundred yards from her office. Later, they also complained about the price of soft drinks in the NIH cafeteria. Was Healy, they asked, skimming off the top to furnish her mansion, "just like [former Stanford University president] Don Kennedy?"

All this was apparently relayed straight-faced back to Healy, who challenged the Dingell staff on the propriety of such 'investigations'. Now, among the subcommittee's other concerns about Healy, Stockton says, "we think she needs to get a sense of humour." **C.A.**

Healy's response to most of this was sympathetic agreement, and reassurances that almost all these concerns were already being addressed. She announced that a two-day director's retreat earlier this month had started the process of a new 'strategic plan' to reform NIH by cutting bureaucracy and streamlining operations. She said her staff were in the process of preparing the first 'site plan' in 20 years, something that could lead to a major new campus — 'NIH North' — to ease the crowding in the Bethesda laboratories. The procurement office is developing an 'action plan' to cut purchasing bureaucracy, she said. And she promised that the NIH's only lawyer — Robert Lahnman — would be getting reinforcements soon.

But she warned the researchers that red tape comes with the territory when one works for the government, especially when an agency has been the focus of a half-dozen congressional investigations in the past year. "Don't shoot the messenger," she said. "We can either break the law and go to jail, or try to improve the law." Presumably more lawyers would help. If not, she wryly joked, "I'm the only person on this campus who can be fired."

Christopher Anderson

Popovic rebuts

Washington

UNABLE to prove that AIDS scientist Robert Gallo is guilty of misconduct in research leading to a blood test for the human immunodeficiency virus, the US National Institutes of Health (NIH) Office of Scientific Integrity (OSI) has turned its fire on Mikulas Popovic, the cell biologist in Gallo's laboratory who was the first to get an AIDS virus to grow in quantity. After the draft of the OSI's report on the case was leaked to the *Chicago Tribune* earlier this month, Popovic's attorneys released a copy of their rebuttal which accuses OSI of following unfair procedures, holding a "clear predisposition" to blaming Popovic for misconduct, and being just plain wrong in some of its interpretations of scientific data.

Once again, NIH have failed to keep confidential documents confidential and, once again, the case is being heard in the press.

Popovic's rebuttal criticizes OSI for withholding access to pertinent documents and blasts the integrity office for using the draft report itself as a vehicle for revealing previously secret information. Thus, Popovic's attorneys Barbara Mishkin and Edward Korwek write: "We now learn for the first time ... that OSI received written answers to interrogatories from Luc Montagnier and the editors of *Science* about the May 1984 publication in *Science* of the Gallo-Popovic paper confirming the viral cause of AIDS. "These answers to interrogatories are pertinent ... yet he was

never told of their existence, much less afforded an opportunity to review them."

Neither Gallo nor Popovic will release the OSI's draft report (in fact, Popovic may sue NIH over the unauthorized leak), but the rebuttal selectively reveals what the draft says. For example, it deals with the question of when and to what extent Gallo and Popovic grew the virus, then called LAV, that came from Montagnier's laboratory at the Pasteur Institute. According to Popovic's rebuttal, Montagnier acknowledged to OSI that as early as December 1983 Popovic told him that he had learned how to "handle" LAV. "Yet the Draft Report nowhere acknowledges this fact. Instead, it rambles for nine pages discussing what Gallo and Popovic may or may not have disclosed to Montagnier, leaving the strong impression that OSI could not confirm what Gallo and Popovic both asserted." Omission of this exculpatory evidence, Mishkin says, is but "one example of OSI's bias".

Then there is the matter of the word "continuous" which has captured the OSI's attention from the beginning of its inquiry. The paper says that Gallo's lab had HIV in continuous culture for five months, during which the culture was reinoculated with fresh cells and virus. OSI says that is not continuous. Popovic says it is, in the ordinary meaning of the word that once virus started growing, it did not stop. The culture always showed virus production.

In reaching a conclusion that Popovic is guilty of scientific misconduct, OSI apparently also takes issue with the reproducibility of his data. According to the OSI draft report, "reproducibility depends fundamentally on good laboratory records and accurate reporting of methods and results." Popovic, acknowledging that his paper could have been written more clearly, nonetheless contends that he has met the test of reproducibility because his experiments have, in real life, been duplicated by scientists around the world.

OSI's central claim against Popovic, based on information currently in the public domain, is that minor errors in the *Science* paper are scientifically "meaningless" but do "gild the lily" by giving the impression that the paper is more thorough than it is. According to Popovic's lawyers, OSI then reasons that gilding the lily constitutes a deliberate intent to deceive. Therefore, Popovic's meaningless errors are intentional. From what the lawyers call a "preposterous resort to baseless psychology", OSI concludes that Popovic has committed fraud.

There is no way for the scientific community to sort this out absent direct access to the many documents in OSI's possession—something it is not likely to get. However, a second OSI report, written with the Gallo and Popovic rebuttals in mind, is now in preparation.

Barbara J. Culliton

End of an era

Washington

SOME 30 years after the US Congress passed landmark legislation decreeing that the only acceptable cancer risk in food was none at all, lawmakers are starting to have second thoughts. Three bills have been introduced this year to overturn US risk-assessment policy, and at least two others are on the way. Spurred by lawsuits from environmentalists and an impending National Academy of Sciences (NAS) report on risk assessment, this Congress may be the one to finally kill the controversial 'Delaney Clause'—the 1958 'zero-risk' provision in US food safety law that many researchers believe has since been rendered obsolete by science.

Last week the Washington-based Institute for Science in Society released a report* comparing the various congressional proposals with current practice and reforms proposed by the White House. The surge of legislative activity, coupled with a growing scientific consensus, "suggests that this Congress may resolve some of the long-standing statutory problems" of US risk assessment, the report concludes. For the researchers, consumer groups and industry associations who have for decades criticized federal risk-assessment policy in general, and the Delaney Clause in particular, such a resolution will come not a moment too soon.

The problem with the Delaney Clause is not that pesticides and food additives are no longer a cancer threat, but that researchers are now too good at spotting carcinogens. Whereas analytical methods for detecting cancer-causing compounds were still primitive enough in 1958 to miss levels that might actually have a health impact on the population, today's methods are sophisticated enough to detect risk levels in the range of one in a thousand million and even lower. In public health terms, such risks are considered negligible. But they are, strictly speaking, non-zero, and a literal reading of the Delaney Clause requires that a food additive or pesticide residue that presents any risk at all be prohibited.

In practice, the Environmental Protection Agency (EPA) and the Food and Drug Administration (FDA) have in recent years chosen to apply a *de minimus* exception to the Delaney Clause, arguing that the spirit of the legislation is to prohibit any *real* risk, a level they defined as greater than 10^{-6} —one in a million.

Hoping to force reforms, however, several environmental and consumer groups have recently sued the agencies on that interpretation, arguing that it is technically in violation of the Delaney Clause and in practice leads to erratic rule-making—FDA and EPA, for example, have often disagreed on the meaning of the

term 'negligible risk'. Perhaps the strongest of the suits, that filed by the Natural Resources Defense Council (NRDC) against the EPA, has won several favourable court rulings and could be decided in NRDC's favour as early as the end of the year.

There is also the issue of just whose risk is to be calculated. Current EPA policy is to base the regulations on the risk to an average consumer, a method that takes into consideration changing consumption patterns over a lifetime. A bill introduced last month by Representative Terry Bruce (Democrat, Illinois) would essentially codify that policy. But Representative Richard Lugar (Republican, Indiana) is expected to introduce another bill that would set a '90th percentile of consumption' level—that of a person who eats more of the particular food than the lower 90 per cent of consumers, but less than the upper 10 per cent. And matching bills introduced in May by Representative Henry Waxman (Democrat, California) and Senator Edward Kennedy (Democrat, Massachusetts) would set the risk levels to that of the most sensitive population group, including infants and people with compromised immune systems.

In general, all the bills would continue to define risks of 10^{-6} as 'negligible' for adults, although the Kennedy-Waxman legislation would set even higher standards when children are at special risk. Critics of the bill say that the practical effect of that provision is to make children the benchmark across the board. "Nobody is going to make food with labels that say 'not for use by children or the immune-compromised'," says Greg Thies, a Senate staff member who is working on the competing Lugar bill.

Nevertheless, thanks to the NRDC lawsuit and the forthcoming NAS study, the Kennedy-Waxman bill is considered the most likely to pass, either this year or in the 102nd Congress's second session next year. If NRDC wins its suit, it will argue for EPA's regulations to be replaced with something akin to the Kennedy-Waxman bill, says senior NRDC attorney Erik Olson.

"If the NAS also concludes that children are at higher risk," says Thies, "I think we'll be looking at a flood of public sentiment and a groundswell of support for the Kennedy-Waxman bill." If so, the Lugar bill (which generally reflects Administration policy) may help to moderate the strict Kennedy-Waxman bill enough to avoid White House opposition.

Christopher Anderson

* *Unraveling Delaney's Paradox: Challenges for the 102nd Congress*, Institute for Science in Society, September 1991.

'Hunting licence' for drugs

Ithaca, New York

A CORNELL University chemical ecologist and a group of Costa Rican conservation biologists have pulled off a novel deal to finance the preservation of biodiversity in that Central American country.

As announced last Friday (20 September), the pharmaceutical giant Merck & Co. has agreed to pay \$1 million over the next two years for the opportunity to screen thousands of unexplored plant and invertebrate species for biologically active molecules that could yield a host of new drugs. The 'hunting licence fee' will go to Costa Rica's Instituto Nacional de Biodiversidad (INBio), a non-profit agency established in 1989 to produce a complete biological inventory for the coun-



A tethered carabid beetle discharging its acid defensive spray on chemical indicator, in response to an 'attack' (pinching leg with forceps).

try — a decade-long, \$50-million effort expected to involve the cataloguing of an estimated 500,000 species, most of which are still unknown to science.

Part of Merck's fee will be passed directly to Costa Rica's ambitious conservation programme, under which more than a quarter of the country's land has been set aside as national parks and reserves.

The rest of the \$1 million will be used by INBio to fund 'chemical prospecting' for biological samples that might contain potentially valuable chemicals. In return, Merck will receive a certain number of samples, as well as exclusive rights to evaluate those samples for pharmaceutical and agricultural applications for a fixed period of time. INBio will receive royalties on sales of any products developed from these samples.

Pamela Adkins, a Merck official, said the company has several products now on the market that were derived from natural sources.

The agreement was conceived by Thomas Eisner, a biologist at Cornell University in Ithaca, New York, and Rodrigo Gámez, INBio's director. The money from Merck will help to train biologists and will pay for field searches for interesting plants and insects. At the same time, with a

\$1-million, three-year grant from the MacArthur Foundation awarded last year, Eisner and Gámez plan to set up a programme to teach Costa Rican scientists the analytical side of modern chemical prospecting: screening for biological activity in a sample using known receptor molecules, then isolating and characterizing the active compounds. The intention is that INBio could eventually use money earned from chemical discoveries made in Costa Rica to finance the conservation of the country's tropical rain and dry forests.

But these financial returns lie many years away. Drugs may spend 20 years in development and testing and, as Eisner points out, a large area of tropical forest can be destroyed in that time. "If you

really want to couple this to conservation, there has to be payment up front," he says. This is where the Merck agreement — which Eisner hopes will be the first of many — comes in.

The first contact with Merck officials was made at a small conference at Cornell in October 1990, and Eisner says he was surprised at how little "conspiratorial thinking" on the part of himself and Gámez was needed to get the company interested in his idea.

In reconciling commerce and conservation, the Merck agreement seems to offer an appealing new way to finance the preservation of biodiversity. But Eisner's plan for Costa Rica has not been unanimously welcomed by his chemical ecologist colleagues. In particular, Timothy Johns, from McGill University in Canada, has dismissed Eisner's approach to conservation in print (*Chemoecology* 1, 142; 1990) as "doomed to failure", because it concentrates on the work of centres such as INBio, rather than involving the local population.

Eisner, who has long been an active campaigner for human rights in Latin America and elsewhere, is sensitive to any suggestion that his plan ignores or exploits local people. But he says that Johns' alternative approach — following up on traditional uses of natural products in developing countries to discover useful chemicals — is simply out of date. Using new receptor-molecule-based chemical screening techniques, many thousands of samples can be screened for biological activity each week, removing the need to rely on such 'ethnobiological' leads. Drug companies have little incentive to invest large sums of money in the sort of programme that Johns would like to develop, Eisner believes.

In any case, Gámez argues that Costa



Blister beetle with droplets of blood emerging from its knee joints. The droplets contain chemical defensive agents, including the well-known substance cantharidin.

Rica's conservation programme does involve local people. The country's national parks and reserves are managed by seven regional bodies, which encourage local people to participate in decision making, he says. INBio is also employing 'parataxonomists' — ordinary people, including high school graduates and housewives, who receive a six-month crash course in taxonomy at INBio before returning to their home towns and villages to work on INBio's inventory in the field. Gámez notes that these people also have an important role as educators in their local communities.

Although Merck's analysis of the Costa Rican samples will initially be based in the United States, Eisner says the training programme that he and Gámez are developing should eventually lead to home-grown screening in Costa Rica. He estimates that 25 drug discoveries are likely over the next decade from Costa Rica's plants and invertebrates — and if Costa Rica eventually receives significant royalty payments from just a few of these, the impact in a country whose entire national budget is less than those of some US universities would be huge.

Nevertheless, Eisner accepts that the Merck agreement might not work in all tropical countries. Costa Rica has a stable, democratic government, and a programme of conservation and cataloguing of its biodiversity that is already fairly advanced — all factors that were attractive to Merck, but absent from many developing countries. Eisner believes that the agreement, if successful, could provide a powerful incentive for other nations to mimic Costa Rica's land use policies in order to attract similar investment.

Peter Aldhous

Stalemate in Nairobi

Washington

Two weeks of global climate talks in Nairobi ended last Friday (20 September) with little substantive progress, as the United States remains faced off against the rest of the developed world on whether to commit to limitations on the emissions of greenhouse gases. Participants reported some movement towards a consensus on the mechanisms for enforcing emissions agreements once those agreements are made, but what the eventual emissions convention will be is still unclear.

The meeting in Nairobi was a precursor to environmental talks next June in Rio de Janeiro, one goal of which is to sign a global agreement on greenhouse gases. With such a covenant nowhere in sight, the delegates decided they will meet in December in Geneva, and may also hold at least one other meeting before next June.

One outcome of last week's talks was an even greater isolation of the United States in the industrialized world, as Japan moved its position closer to that of the European Communities (EC). The EC wants industrialized countries to commit to holding their total emissions of carbon dioxide at 1990 levels after the year 2000. Japan, which previously had not agreed that such commitments should be part of a global climate convention, did put its weight behind commitments, but with a number of qualifiers. Industrial countries should, Japan proposed, make a "best effort" to stabilize emissions of carbon dioxide and other greenhouse gases "as soon as possible, for example by the year 2000, in general at a 1990 level, recognizing the differences in approach and in starting point in the formulation of objectives".

The United States, according to the policy of President George Bush, does not believe that the scientific evidence for global warming is solid enough to commit to limiting greenhouse gases. "It simply

ARCHAEOLOGY

Dead Sea Scrolls released to scholars

THE Huntington Library in the Los Angeles suburb of San Marino, California, announced last week that it will release to all qualified scholars photographs of the Dead Sea Scrolls. The library has some 3000 photographs of the scrolls and claims that it is under no legal obligation to keep them secret, although its position may be challenged by scholars who have been working with the original scraps of parchment since the scrolls were discovered in 1947. The move by the Huntington follows publication three weeks ago of an unauthorized edition of the scrolls produced with the aid of a concordance and a computer (*Nature* 353, 96; 12 September 1991). B.J.C.

does not seem to us to make practical sense [to make commitments for emissions cuts] at this time in light of the uncertainties," said Robert Reinstein, a deputy secretary of state who led the US delegation to the talks, in a prepared statement. Furthermore, the US position holds that with the phase-out of chlorofluorocarbons (CFCs) by 2000, as agreed to under the Montreal Protocol, and with the provisions of the 1990 Clean Air Act and Bush's 'national energy strategy', which calls for increased energy efficiency and more nuclear power, the United States will keep its emissions of greenhouse gases effectively stabilized at 1990 levels from 2000 to 2030.

During the talks, the Japanese cleared up most of the confusion about their proposal of 'pledge and review', which they offered at this past summer's environmental talks in Geneva. At the time, the concept was attacked by environmentalist groups as a way to avoid making solid commitments (see *Nature* 352, 3; 4 July 1991). In the Nairobi talks, Japan and the EC both agreed that 'pledge and review' would be a method of carrying out commitments previously made — a 'pledge' would be a clearly delineated programme for meeting the commitments, which would be reviewed periodically to see that it was being met, although there is no agreement on who would do the reviews.

The delegates succeeded in putting together a set of 'working papers' that will be consolidated and used as a starting point for negotiations in December. That package, however, is no more than a 'wish list' with mutually contradictory proposals from the various delegations. Meeting participants say that the compromises made over this list will be the hard work of the next conference and will determine just what kind of global climate treaty will be possible in June.

There is broad enough consensus on certain issues that it seems likely that some agreement will be reached in time for the meeting in Rio. But these issues are the less challenging ones, such as the structure of the institutions that will oversee an eventual climate treaty and the question of technology transfer and financial aid for developing countries.

At present, given the US position, it seems certain that next year's climate convention will not contain specific commitments for greenhouse gas emissions, at least not for all industrialized countries. Some countries, however, have indicated that they will not agree to a treaty unless it has such commitments. The trick then will be to get enough consensus on other issues that some substantive covenant can be signed as a starting point for future agreements as needed. **Robert Pool**

Another major deal

Washington

WALL Street and the US biotechnology community were once again taken by surprise last week when American Home Products said it has agreed to acquire a 60 per cent stake in the Massachusetts-based Genetics Institute, in a transaction valued at \$666 million. The announcement comes hard on the heels of the decision two months ago by California-based Chiron to acquire neighbouring biotechnology company Cetus (*Nature* 352, 364; 1 August 1991). News of this latest deal serves to underscore the research- and capital-intensive nature of the industry.

The agreement, which shares many similarities with last year's marriage of Genentech to the Swiss health-care conglomerate F. Hoffman-La Roche, provides Genetics Institute with \$300 million in cash — money that should enable the company to accelerate drug development. At the same time, it allows American Home Products, a New York-based drug and consumer goods company with sales of just under \$7,000 million in 1990, to establish a strong foothold in biotechnology.

Genetics Institute has taken some knocks with its first two products, tissue plasminogen activator (TPA) and erythropoietin (EPO). In May 1990, under a licence agreement with Genetics Institute, Wellcome abandoned a six-year research effort to develop TPA (*Nature* 345, 194; 1990). But perhaps more devastating for Genetics Institute was a decision last March by the US Court of Appeals that overturned a lower court's decision and dismissed the company's US patent claims to EPO, a drug that stimulates the body's production of red blood cells. The decision left the rival company Amgen with a lucrative market monopoly for its EPO product in the United States (*Nature* 350, 99; 14 March 1991). A decision by the US Supreme Court on whether to review the appeals court's ruling is expected this autumn.

Although Genetics Institute has just turned the corner to profitability, Gabriel Schmergel, president and chief executive officer of the company, says he was still faced with the prospect of having either to cut or to postpone key projects. The \$300 million in cash "is a sum of money that we could not have raised any other way", says Schmergel, who was approached by American Home last spring. Over the next two years, Schmergel says he plans to increase the current 600-member staff with 125 to 175 new hires. He also says the additional resources will enable Genetics Institute to keep development of its three earlier-stage products in-house. These include a bone-inducing protein, BMP-2; the platelet growth factor interleukin-11; and the blood-cell growth factor macrophage colony-stimulating factor. Diane Gershon

A \$500-million renovation

Tokyo

In a move to infuse some new life into Japan's rundown national universities, the Ministry of Education, Science and Culture is hoping to get more than \$500 million over the next five fiscal years to renovate and rebuild them. The rebuilding plan is the most significant attempt in years by the ministry to repair Japan's decayed university infrastructure in the face of growing criticism.

The huge request, which starts with a budget of more than ¥15,000 million (\$111 million) for the next fiscal year and continues with similar amounts in subsequent years until fiscal 1996, comes on top of substantial requests by the ministry for increased funds for grants and fellowships to support university research (*Nature* 353, 102; 12 September 1991). The request is subject to approval by the Ministry of Finance and the Diet, but although some trimming may occur, substantial changes are unlikely.

A ministry official in charge of the renovation programme says the ministry decided to make the move partly because of savage press treatment of the shoddy state of Japan's universities. But in recent years, industry and the powerful Ministry of International Trade and Industry have also become increasingly vocal in expressing their concerns about the universities. Industry fears that there will soon be a dearth of young scientists and engineers in Japan to support industry, partly because of a rapid demographic decrease in the student-age population, but also because of the unappealing nature of Japan's rundown universities. With even cleaning staff and technicians in short supply, new buildings alone are no longer enough.

Probably the most vocal (or at least the most powerful) critic has been Akito Arima, president of Tokyo University. And although education ministry officials emphasize that it has not yet been decided how the money would be spent, the science and engineering faculties of Tokyo University, which are undergoing major reform (*Nature* 351, 679; 27 June 1991), have the most concrete plans for renovation among Japan's many national universities.

The engineering faculty already has a plot of land set aside on the main campus of Tokyo University for a planned new ten-storey building. Administration officials hope to begin construction next fiscal year with some of the funds requested by the education ministry.

The science faculty is in a more difficult situation because it does not have any spare land. If the main faculty building is torn down, there would be nowhere to accommodate the present faculty mem-

bers, supporting staff and students, who total about 300 people. Present plans call for half of the main building to be torn down and replaced with a much taller, ten- to fifteen-storey building, and then the other half would be replaced. Rather than just holding the physics department as at present, the new building would accommodate all the science departments except biological sciences and chemistry, which already have their own new buildings on campus.

BRITISH RESEARCH SPENDING

SERC ponders partnerships

London

BRITAIN'S biggest research council, the Science and Engineering Research Council (SERC), seems reconciled to a further period of budgetary constraint. But SERC's chairman since 1 April, Sir Mark Richmond, seems to relish the role of good housekeeper now thrust on him. Most big projects and even the council's own decision-making structure are due for review, to be completed by early 1993.

This emerges from SERC's 'corporate plan', published last week. One issue, a problem which has unexpectedly become an opportunity, is Europe's over-supply of neutrons for research. With the enforced hiatus in the operation of the nuclear reactor at the Institut Laue-Langevin at Grenoble (*Nature* 352, 366; 1 August 1991), researchers are beating a path to SERC's Neutron Spallation Source at the Rutherford Appleton Laboratory.

Richmond boasted last week that SERC's neutron source has already become an international centre, to the cost of which a handful of other research councils (Japanese as well as European) already contribute. But he also hinted that continued British membership of the Grenoble neutron laboratory would hang on the estimated cost of fixing the Grenoble reactor. The upper limit of £100 million overall would be too much.

Another bout of introspection about continued British participation in the European high-energy physics laboratory (CERN) at Geneva is also on the cards. SERC's Nuclear Physics Board has enthusiastically endorsed the plan to install in the tunnel now housing LEP (the Large Electron-Positron accelerator) a pair of accelerators for protons and antiprotons that will make the Large Hadron Collider.

Richmond is content that the construction costs will fit within the present budget (so that the British annual contribution should suffice), but is alarmed that the high cost of detectors will strain the discretionary budget, perhaps restricting Brit-

Given the complications, the science faculty does not expect to begin reconstruction for two years. And even meeting this schedule will be difficult, faculty members say.

The ministry's plans to put more money into the universities in the form of grants, fellowships and new buildings goes part of the way towards meeting the widespread criticisms of the university research system. It remains to be seen, however, how much other universities apart from Tokyo University will benefit from the rebuilding programme.

David Swinbanks

ish participation at CERN. SERC as a whole will have a first discussion of the project later this year; CERN's management wants a final decision by the end of 1992. Richmond, an enthusiast for European projects, hopes that a planned reassessment of national contributions to CERN or a possible delay of installation costs will liberate SERC from the financing jam.

That is likely to be acute. This year's budget of £454.9 million is 4.0 per cent less than last year's in real terms, even after using the Treasury's estimates of inflation in the general economy, but the costs of research increase more quickly. And SERC is faced with budgets that will decline (on the Treasury's basis) by 2.3 per cent next year and 4.0 per cent in 1993-94. One consequence is that SERC's staff will be cut by 300 people; some redundancies will be involuntary.

Two aspects of the review of SERC's own structure are especially significant. First, Richmond is looking for a system in which the four disciplinary boards (nuclear physics, astronomy and space, science, and engineering) that at present compete for shares in SERC's pie may be restructured to reflect the growth of interdisciplinary science. In a manner likely to win friends among politicians, he is also looking for better ways of making links between academic research and industry more productive of useful innovation.

The question (raised last week) of whether responsibility for engineering should be generally distributed through SERC's activities rather than concentrated in an engineering board (due to spend £138 million next year) is less likely, when asked aloud, to win friends among engineers. SERC's spending on engineering has had a talismanic quality since the research council added the 'E' to its name in 1982: Richmond's safest course may be to ensure that the promised detailed reviews of spending on engineering are widely publicized.

John Maddox

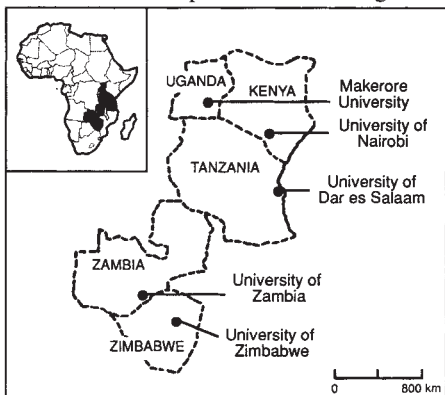
Opening up communications

Washington

FOR medical researchers and doctors in the developing world, one of the greatest handicaps is the lack of communication with their peers. The medical library at Uganda's Makerere University, for example, has had no subscriptions to Western medical literature for more than a decade — it cannot afford journals that require payment in foreign currency. And contact with other countries by phone or fax is generally unreliable and sometimes almost prohibitively expensive.

But starting in October, health workers in five African universities will have an alternative, thanks to a recently launched 45-kg satellite and to SatelLife, a non-profit organization based in Cambridge, Massachusetts.

SatelLife is establishing a communications system that will transmit medical literature, including *The New England Journal of Medicine*, to libraries in developing countries and will allow researchers and doctors in those countries to communicate with their peers around the globe.



The system will begin with ground stations in five locations in Africa: Makerere University, the University of Dar es Salaam in Tanzania, the University of Zambia, the University of Nairobi and the University of Zimbabwe. Within 12 months there should be 20 stations in Africa, says Charles Clements, executive director of SatelLife.

Each \$5,000 ground station consists of a personal computer linked to a radio that can communicate with a satellite launched in July, whose polar orbit takes it around the Earth once every 100 minutes. When it passes over one of the ground stations, it automatically transmits any messages addressed to that station and uploads files waiting in the station's computer. The satellite stores these message files and later beams them down to the other stations to which they are addressed. This 'store and forward' system allows stations to pass messages back and forth. Although the transmissions are not immediate unless two stations are both within range of the satellite at the same time, the satellite passes

over each spot on Earth at least twice a day, so messages are delivered quickly.

A ground station in Newfoundland will send electronic copies of several medical journals to the satellite for transmission and will also allow people on the system to communicate with researchers in North America and Europe through electronic mail connections. The Harvard School of Public Health has agreed to provide consultation on clinical and public health matters through the satellite link.

Firoze Manji of the International Development Research Centre in Nairobi said that although access to Western literature and researchers is important, the most valuable aspect of the satellite system will be the ability of researchers and clinicians from developing countries to communicate with each other. These workers share similar problems and interests, but they seldom know what each other is doing because of unreliable, expensive telephone systems. Sending a six-page fax from

Benin to Nairobi can cost \$300, Manji said, and it is often not even possible to call or send faxes between East Africa and West Africa. "This is one of the single most important initiatives to develop research capabilities in the Third World that has been undertaken since the Second World War," Manji said.

The greatest obstacle to the spread of this communications system may be overcoming government resistance, said Marvellous Mhloyi of the University of Zimbabwe. Some postal and telephone authorities may not want to give up control over and profits from communications in their countries, and some governments may be suspicious of the system because it originated in the West.

Eventually, SatelLife hopes to expand its system across Africa and into other developing countries. The existing satellite can handle 500 ground stations, and a second satellite is planned for 1993 or 1994, which will double both the capacity and the number of times a day that a satellite passes over each ground station.

Robert Pool

NEWS IN BRIEF

Bad loans, good science

Washington

ALTHOUGH the concept is still new, 'debt-for-nature' swaps have helped preserve more than \$60 million of land in developing countries over the past three years. That may not mean much to the Brazilian rain forest, but the relative success of the initiatives have not escaped the notice of US Representative George Brown (Democrat, California), chairman of the House science committee. If such deals can work for conservation, he asks in an upcoming article in *Issues in Science and Technology*, why not research?

Brown inserted language into last year's budget authorization directing the Agency for International Development to spend up to \$20 million on debt-for-science aid for Mexico. But to date, the State Department has simply ignored the request, choosing to focus on higher aid priorities in Eastern Europe and the Middle East instead.

Now Brown is trying again. Last month he introduced a bill that would authorize the National Science Foundation (NSF) to spend \$5 million on Latin American debt-for-science trades. If passed, the bill would instruct NSF to make the money available to universities so that they can buy up foreign debt, which often trades at far less than its face value thanks to the poor repayment histories of many of the Latin American countries.

In return for having its red ink erased, the recipient nation would promise to spend at least 75 per cent of the debt's face value or 20 per cent more than its going price on cooperative research in agriculture, health

and the environment. Given the late introduction of the bill, no action is expected until the 1993 budget comes up for debate next year.

C.A.

Trials by wire

Washington

FOR clinical researchers with enough computer hardware on their desks, the future of scientific publishing is now. This week the American Association for the Advancement of Science (AAAS) launched the first all-electronic peer-reviewed medical journal, *Current Clinical Trials*. Using custom graphics software (based on Microsoft Windows 3.0, and inheriting that package's substantial memory, monitor and processor requirements), a subscriber will be able to view typeset-quality text and figures, see new articles within 24 hours of their acceptance, and search for keywords, authors and data of interest.

Owners of more humble personal computers can receive a text-only version of the journal, and subscribers can arrange to be alerted by fax every time an article on AIDS trials, for example, is published. Readers will be able to view the content of any reference cited in the MEDLINE online database by clicking on the reference, and in future releases will be able to get full definitions of methodological terms the same way. Counting on an all-electronic peer-review and production process to eliminate many of the delays of tradition publishing, AAAS has set an optimal turnaround time of fewer than three weeks between submission and publication. A yearly subscription to the journal will cost \$110.

C.A.

Energy costs of a long life

SIR — The publication in 1989 of a stamp commemorating the centenary of the German social security system (see figure) has, probably inadvertently, focused attention on the widely neglected phenomenon of the virtual doubling of the average life expectancy in advanced industrial societies over the past 100 years. Whereas the age distribution of agrarian or early industrial societies displays the shape of a pyramid typical of most biological populations (left-hand graph), the demographic picture of developed communities is reflected by a balloon-shaped function, the narrow base of which constitutes an artefact caused by the introduction of oral contraceptives during the 1960s (middle and right-hand graphs in the stamp).

Although many will ascribe this demographic miracle to a multitude of specific causes (notably the advances of biomedical sciences during the past century), the root cause deserves to be stated explicitly. The principal factors promoting longevity in humans (balanced nutrition, evasion of excessive toil, provision of medical care, social security and so on) are corollaries of developed societies characterized by sizeable gross national products (GNP). Because of the well-established correlation between GNP and energy consumption, it seems evident that the conspicuous doubling of the average life-span of whole populations (not of a minority) over the past 100 years is linked to the progressively increasing energy flow through society over the same time. Biological analogues demonstrate that similarly conspicuous innovations had been generally dependent on preceding discrete quantum steps in bioenergetic evolution involving upsurges in physiological energy consumption by about an order of magnitude and more¹. This is approximately the factor by which the industrial ('extrasomatic') energy flux surpasses the per capita somatic flux in modern industrial societies (for instance, the 1.6-million population of the city of Hamburg has a physiological energy consumption of some 5,000 TJ yr⁻¹, whereas the energy expenditure of the community as a whole totals 100,000 TJ yr⁻¹).

A glaring demonstration of the interrelationship between the energy-related GNP and life expectancy has been furnished by the temporarily diverging socio-economic orders of former East and West Germany, with average life spans cut by 2.5 and 7 years for men and women, respectively, from the East². Specifically as a result of this historical

experiment, it is clear that average life expectancy for broadly the same human gene pool is apt to vary significantly in response to the socio-economic environment, with a conspicuous advantage for the more voracious energy consumers.

With these interrelationships established, it is not difficult to predict that there will be no voluntary way back from a high-energy to a low-energy economy. Naïve criticism of modern industrial society ignores the privileged condition of its members in terms of both life expectancy and general diffusion of affluence, disregarding the fact that the recent rise of energy consumption by *Homo sapiens* to



The 1989 stamp compares age composition of an agrarian/early industrial society (1889) to that of an advanced industrial society (1989 and projection for 2000). Note that the population 'pyramid' typical of poorly developed (low-energy) societies assumes the shape of an elongated balloon for high-energy societies (in these examples, this shape is gravely impaired by scars left by two major wars).

levels 10–20 times above the basic metabolic rate is glaringly beneficial to the species.

Although the increased energy flow through the human ecosystem is evidently responsible for this most important single improvement of the human condition in the history of mankind, its intrinsic potential for the emergence of an autotoxic (self-poisoning) effect due to the accumulation of toxic waste products of the industrial process remains a matter of concern³. The future of high-energy (industrial) societies will, accordingly, depend on man's capability to maintain an adequate energy flow while minimizing the adverse effects of his extrasomatic metabolism on the environment by suitable technological fixes⁴.

MANFRED SCHIDLowski

Max-Planck-Institut für Chemie,
Postfach 3060,
6500 Mainz,
Germany

1. Wieser, W. *Naturwissenschaften* **73**, 543–549 (1976).
2. Casper, W. & Hermann, S. *Ing. Bull. Inst. Med. Statistik u. Datenverarb. Berlin* **13**, 9 pp. (1990).
3. Rowland, F. S. & Isaksen, I. S. A. (eds) *The Changing Atmosphere* (Wiley, Chichester, 1988).
4. Schidlowski, M. *Terra Nova* **3** (in the press).

Smart whale

SIR — Daedalus (*Nature* **352**, 384; 1991) described a novel 'heavy-waterglider' whose undulating trajectory derives from the depth-related changes in density that follow from the unusual thermal properties of heavy water. Fact is even better than fantasy, for his design has been largely pre-empted by the sperm whale. The substitution of spermaceti oil for heavy water, and the closer control of heating and cooling by changes in the internal blood supply, make the sperm whale a much more flexible undulator¹. It also has an operationally tried and tested 'smart' navigation system, instead of an automatic one. Its celestial navigation system is still classified information, the details of which are transferred only on a 'need to know' basis for regular migrations. Daedalus's vehicle "will be guided to rendezvous . . . by sonar steering commands". The sperm whale's mastery of the language of sonar completely outclasses its vehicular competitor, but the rendezvous of its choice is more likely to be a mate or a family party that Daedalus's "receiving tug".

As part of the promotional strategy for the vehicle, Daedalus asserts that "oceanographers will love the heavy waterglider". He should know that we already do love the sperm whale.

PETER J. HERRING

Institute of Oceanographic Sciences,
Deacon Laboratory,
Brook Road,
Wormley,
Godalming,
Surrey GU8 5UB, UK

1. Clarke, M. R. *Nature* **228**, 873–874 (1970); *J. mar. biol. Ass.* **58**, 27–73 (1978).

Grand spectacles

SIR — As a sufferer from presbyopia, I have longed for the 'adaptive optics' spectacles proposed by David Jones (Daedalus, *Nature* **352**, 287; 1991). I have two suggestions for improvements.

First, the eye has significant chromatic aberration, so the infrared image will have to be slightly out of focus in order to focus the optical image perfectly. Second, I wish to retain what focusing ability I still have for occasions when I do not use my spectacles. I suggest therefore a bias and lag in the automatic adjustment, calculated to exercise my eye muscles so that they do not atrophy. Both effects would be simple for the microprocessor controlling the optics.

NELSON MAX

Lawrence Livermore
National Laboratory,
Post Office Box 808,
Livermore,
California 94550, USA

In the volcano's shadow

Daisuke Shimozuru

Earlier this year, Unzen volcano erupted, killing 42 people. What follows is an account of how Japan's volcano coordinating committee dealt with this and similar events.

IN a recent News story (*Nature* 351, 511; 1991), David Swinbanks reported that Japan's coordinating committee for the prediction of volcanic eruptions had failed to predict three recent episodes: the pyroclastic flow of Unzen volcano on 3 June 1991; the fissure eruption of Izu-Oshima volcano on 21 November 1986; and the small submarine eruption off Izu peninsula on 13 July 1989. Swinbanks does not report events completely accurately and, on behalf of the committee, I would like to explain exactly how we coped at the time.

Unzen volcano

Seismic activity began in November 1989 at Chiji-iwa Bay, about 13 km west of Fugen-dake (the erupted cone of Unzen volcano), and gradually migrated eastwards from July 1990. The Meteorological Agency despatched an observation team in mid-October to intensify seismic monitoring, assisting the agency's local observatory. Early in November, volcanic tremors were observed. At that time, we anticipated the impending eruption and asked the Ministry of Education, Science and Culture on 9 November for financial support for more seismic monitoring by university groups. That very day, the ministry provided the money. The university team set up a seismic net in cooperation with the Shimabara Earthquake and Volcano Observatory (SEVO) of Kyushu University. SEVO is in Shimabara city, at the east-northeast foot of the volcano; later it became the base of the telemetered data and advance headquarters of the committee.

On the morning of 17 November, a minor phreatic eruption occurred in the summit crater of Fugen-dake. We could not predict the date and time of the first outbreak. Strangely, no shallow volcanic earthquakes were observed beneath the erupted centre. After this first outbreak, serial volcano information was issued from the Meteorological Agency's local observatory, calling the public's attention to the situation. Furthermore, the possibility of debris flows triggered by rain were included in this information because of the ash deposits due to repeated phreatic eruptions.

From 13 May, extremely shallow earthquakes began to occur just beneath the eruptive crater. We expected a lava extrusion and issued an official committee chairman's comment. Three days

after this comment, lava was first seen in the crater. On the morning of 24 May, the first small-scale Merapi-type pyroclastic flow occurred, descending to the small valley of the eastern slope of the volcano. The first debris flow occurred on 15 May, followed frequently by others. The local government recommended that people living along the river leave the danger zone at the time of heavy rains. On 26 May, a worker got slightly burnt by the hot ash cloud of the pyroclastic flow. Taking account of this event, an evacuation order was issued by the local government on 26 May for the danger zone, and a restricted area was clearly marked. The Meteorological Agency's local observatory and the committee warned, also on 26 May, of future pyroclastic flows as well as debris flows.

On 31 May, a committee meeting was held in Tokyo, where evaluation of the

current activity of Unzen volcano was discussed. An official statement was issued to the public, warning of pyroclastic flows. The chairman proposed to set up the advance headquarters of the committee at SEVO to be able to respond to rapid change in volcanic activity. Kosuke Kamo was appointed deputy chairman of the committee. He arrived at SEVO on 4 June, one day after the occurrence of tragic pyroclastic flow that claimed 42 lives, including those of Maurice and Katia Krafft and Harry Glicken.

The local government urgently needed a scientific basis to set up a restricted zone, and asked the experts on 1 June to produce a hazard map of pyroclastic and debris flow. They prepared a simulation map by 4 June. Based on this map, the local government created an exclusion zone, with the result that 12,000 people were evacuated from the area by 10 June.

Concerning the 3 June tragedy, I would like to stress the following three points. (1) The time of occurrence and magnitude of discrete pyroclastic flows is hard to predict. We tried to find any characteristic feature of the precursors, but found none. (2) The appearance of pyroclastic flows, in particular growing hot ash clouds and fast descent, were frequently televised. Accordingly, I guess that at least 10 million people might have been aware of their violence. The injured worker appeared on television on 26 May. (3) Before the 3 June tragedy, the local government had already defined the exclusion zone along the river. All residents had already been evacuated from this area. (4) To obtain spectacular shots of pyroclastic flows, many television and press staff, including the three volcanologists I mentioned above, were inside the restricted area. They were killed by hot ash clouds.

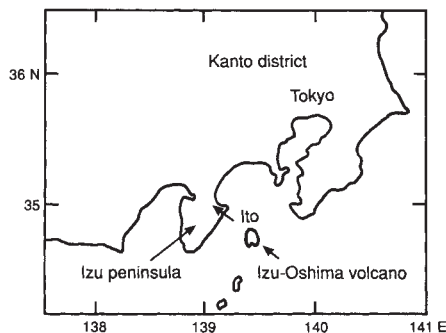
Izu peninsula

The Eastern part of Izu peninsula is characterized by a monogenetic volcanic field in which more than 70 cones are widely scattered, even off-shore. The eruptive event occurred about 2,000 years ago inside the peninsula. The South Kanto district, including Izu peninsula, has been the designated area for intense geophysical monitoring by the Committee for Earthquake Prediction. Actually, earthquake swarms,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

A fireman and fire engine escaping along a road in Kitakoba community, Shimabara in southwest Japan as a massive flow of lava, molten rocks and ash erupted from Mount Unzen in June. (AP)

although the duration period is short, have been observed since 1978 at the east coast of Izu peninsula as a yearly event. In 1989, familiar earthquakes began to occur from 30 June. On 5 July, a subcommittee of the Earthquake Prediction Committee met to evaluate this activity. On 9 July, an earthquake of magnitude 5.5 took place, followed by a sharp decline of seismic activity. Discoloured sea water at several spots near the coast was reported by local people on 5, 7, 8 and 11 July. On reflection, most of the scientists concerned thought the observed seismic and deformation data resulted from shallow-seated magmatic action; however, we were unaware of the critical situation.



Suddenly, from the evening of 11 July, after the seismic activity declined, a severe tremor started. At that moment, I rushed to the Meteorological Agency and held a press conference at midnight. I reported my personal opinion, that a bubbling or minor steam explosion had started at the sea bottom. The next day, I convened the executive board of the committee and issued a chairman's comment, that the tremor had possibly been generated by magmatic action beneath the sea bottom. We added that there were no definite surface phenomena indicating eruption at that time, and stressed the necessity of further careful observation. The epicentral area was very close to the popular resort of Ito city, and we were afraid that if we suggested the possibility of an eruption, we would cause mass panic.

To our regret, we could not clearly predict the impending eruption. The fault activated by the earthquake was estimated to have an east-west direction and extend to the northern part of Ito city. We thought that magma would extrude through a certain spot along this fault. Tilt and electromagnetic distance measurement (EDM) data, as well as a volumetric strain meter, showed marked changes associated with the seismic activities, but no precursory data were found before the eruption. We could not predict where and when the eruption would take place. At last, on 13 July the submarine eruption was observed between 1830 and 1845, 3.6 km northeast off the coast of Ito city. This was our

first experience of attempting to predict a submarine eruption. The series of seismic and eruptive episodes posed a problem for the relationship and cooperation between the committees for Earthquake Prediction and Prediction of Volcanic Eruptions.

Izu-Oshima volcano

There has been a gradual decrease in total magnetic intensity since 1983 at Oshima volcano. The Izu-Oshima Volcano Observatory of the Earthquake Research Institute of the University of Tokyo was established in 1985, and monitoring has been taking place since then. Before the completion of this observation system, intermittent volcanic tremors began to occur from July 1986. Simultaneously, apparent electrical resistivity beneath Mihara-yama, the central cinder cone inside the caldera, showed a remarkable decrease. We intensified our observations, including levelling, EDM, gas and temperature of the crater floor. Most of the data suggested that there was increasing magmatic activity beneath Mihara-yama. However, continued subsidence of Mihara-yama relative to the caldera rim has been observed by repeated levelling.

A regular meeting of the committee was scheduled in October 1986, but in view of the observed data, we could not wait. Therefore, I convened a meeting of the executive board of the committee on 14 August and on 24 September to evaluate the data. Earthquake activity was not apparent on the island. Our final conclusion was influenced by the continuing subsidence of the central cone, though most of the abnormal data were concentrated beneath it. Previous investigation suggested the existence of a magma reservoir 3–4 km beneath Mihara-yama. Inflation of Mihara-yama could have been expected if a large volume of magma was supplied beneath it. As a result, the committee warned of the possibility of an eruption from the summit crater of Mihara-yama, but there were no data to suggest that a large eruption would occur elsewhere on the island.

On 12 November, steam was observed issuing from the inside wall of the summit crater of Mihara-yama. On 14 November, an executive board meeting was again convened. We were convinced that there would be an eruption from Mihara-yama soon, but could not predict when. The next day, the eruption started inside the crater. Strombolian-type eruptions continued for almost one week, and the lava overflowed the crater rim in the morning of 19 November. In the morning of 21 November, there was an apparent lull. Then, severe earthquakes began around 1400 h in the northwest outside the caldera. In the meantime, a

fissure opened on the caldera floor at the northwest foot of Mihara-yama at 1615 h. One-and-a-half hours after this, a new series of vents opened outside the caldera rim. We did not have enough time (only 2 hours) to interpret the sudden occurrence of severe earthquake swarms. To our regret, we could not predict the fissure eruptions inside and outside the caldera. The last fissure eruption from outside the caldera was 565 years earlier, in 1421, in the south-western part of the island.

Conclusions

During the past 6 years, Japan has experienced three different types of eruption: the fissure eruption at Izu-Oshima volcano in 1986; the submarine eruption off the coast of Ito city in 1989; and the 1991 eruption of Unzen volcano accompanied by the dome growth and pyroclastic and debris flows. Thanks to the National Programme for Prediction of Volcanic Eruptions, which started in 1973, volcano observatories have been equipped with sophisticated and modern instruments. We can detect any data suggesting impending eruptions. But prediction of eruptions is extremely difficult. Even when we are convinced that the magma system beneath a volcano is in a critical state, sometimes the extrusive event does not occur – this is a stillborn eruption. On the other hand, eruptions are sometimes triggered by external disturbance such as rainfall or lunar-solar tidal force. Consequently, prediction of the date and time of initial outburst is still like a kind of betting.

Basically, eruptions are caused by fracture of solid rock and viscous magma, which is governed by a stochastic process. This means eruption prediction could only be done in a probabilistic manner, which would have fatal consequences. Unfortunately, there are too few precursory data for volcanoes that only seldom erupt. Another difficult problem is to predict objectively the future of ongoing eruptive activity and, consequently, to give suggestions to government officials for the timing of removal of evacuation orders.

Recently, the Commission on Mitigation of Volcanic Disasters, via the International Association of Volcanology and Chemistry of the Earth's Interior, proposed the "Decade Volcano Project", as part of the International Decade for Natural Disaster Reduction. Such international cooperation will doubtless promote eruption prediction and hazard mitigation. □

Daisuke Shimozuru is chairman of the Coordinating Committee for Prediction of Volcanic Eruptions, Kyodo 1-17-7, Setagayaku, Tokyo 156, Japan.

■ See the letters on page 307 of this issue.

AIDS research turned upside down

Two recent developments support the view that HIV is a necessary but not sufficient cause of AIDS: if confirmed, they must cause a benign revolution in the search for better treatment and for prophylaxis.

PROFESSOR Peter Duesberg from the University of California at Berkeley is probably sleeping more easily at night now than for five years, since he first took up cudgels against the doctrine that AIDS is caused by the retrovirus HIV (or human immunodeficiency virus). Duesberg has been pilloried for his heterodox views (best put in *Proc. US Nat. Acad. Sci. U.S.A.* **88**, 1575; February 1991), and faced with the threat that his research funds would be snatched away. Now there some evidence to support his long fight against the establishment (among which, sadly, he counts this journal).

The evidence comes in two mutually consistent parts. Earlier this month, Tracy A. Kion and Geoffrey W. Hoffmann from the University of British Columbia published an account of experiments in which mice that had been treated with T lymphocytes from another mouse strain, but not exposed to HIV in any form, had been found to generate antibodies against two of the proteins characteristic of HIV — those known as gp120 and gp24 (*Science* **253**, 1138; 6 September 1991).

But how can a mouse develop antibodies against protein constituents of a virus of which it is apparently innocent? The authors' explanation is embodied in a theory put forward before the experiments were done (Hoffmann *et al.*, *Proc. Natn. Acad. Sci. U.S.A.* **88**, 3060; April 1991). Briefly, it is that AIDS is an essentially auto-immune disease in which T cells have lost the normal 'anti-self' interdiction, but instead kill each other. The viral coat-protein gp120 is concerned because, Kion and Hoffmann argue, it is likely to bear a close structural similarity to class-II MHC proteins (where MHC means 'major histocompatibility complex').

The other piece of evidence (to be published as Scientific Correspondence in *Nature* next week) is an account of observations on the protection of macaque monkeys against the simian analogue of HIV called SIV (simian immunodeficiency virus) by means of a vaccine consisting of monkey T cells first infected with SIV and then inactivated. The group concerned is at the National Institute for Biological Standards and Control in Hertfordshire. E. J. Stott and eight colleagues describe how they immunized a group of four macaques with their inactivated cellular vaccine, and four others with identical cells that had not been infected with SIV.

That three animals in the first group

were protected against later challenge with SIV would have been good news for the vaccine-makers. The first surprise is that two out of the control group of four were also protected. And the sensation is that the essential difference between the protected and unprotected monkeys is that the former have ten times as much antibody against components of the T cells used as virus carriers as do the latter — the control group susceptible to SIV infection on later challenge (see story page 287).

Again, it seems, the presence (or absence) of SIV is not the sole determinant of the production of antibodies apparently diagnostic of vulnerability to infection. (In these experiments, antibodies against the protein components of SIV have been looked for, but have not been found.) Hunting through past records has allowed the group to extend its suspicions to vaccines based on purified (and then inactivated) SIV particles as well as to another culture of T cells; again it seems (with 40 animals) that the distinctive diagnostic of protection against later challenge with SIV is the level of antibody against the T cells from which the vaccine was derived.

It goes without saying that the British results would have been even more persuasive that something strange is afoot if the vaccination experiments had more often been accompanied by adequate control groups. It is not so much that monkeys are scarce, but that the general inhibition of their use has persuaded people that corners should be cut. The result may well be that an important qualification — to say the least of it — of the standard picture of AIDS infection may have been needlessly delayed.

The idea that the mechanism of AIDS may be that of an auto-immune disease helps to explain some of the special features of AIDS, of which the usually long interval between infection (with HIV) and the emergence of the eventually fatal symptoms are two. Another is Duesberg's main plank in his campaign against the conventional wisdom: that virus can be recovered from only a small proportion, perhaps 1 in 500, of the T cells in a person infected with HIV; how, he asks, can that be reconciled with the view that the depletion of T cells is caused by the multiplication of virus in infected cells and their spread to others?

Hoffman's theory is more specific. The starting point is the mechanism of self-tolerance in which T cells are selected in the thymus not to react strongly against an

organism's own MHC molecules. What happens, then, when foreign lymphocytes find their way into a person's blood? The immune system will make antibodies against the recognizable antigens on the surfaces of the cells, and T cells will be organized to kill off the intruders. But then suppose that HIV is simultaneously injected into the blood; the immune system will similarly dispose itself against that foreign material. But if the cunning of HIV is that gp120 closely resembles the standard structure of class-II MHC molecules, there will then be two armies of T cells organized to attack each other.

That, at least, is a simplified form of what Hoffmann says. In reality, there are several different kinds of T cells, the interaction between them is a dynamic process and the most that can be said is that, after a simultaneous exposure of a person's immune system to both foreign lymphocytes and HIV, the usually delicate dynamic balance between cells with different functions and specificity is undermined.

It would be wrong to suppose that this is what really happens in the course of AIDS. (Recent studies suggest that gp120 and MHC class II bind to different sites on the CD4 molecule, the receptor for gp120: Fleury *et al.*, *Cell* **66**, 1037; 1991.) But one of Hoffmann's more recent observations may carry weight; mice bred for auto-immune disease (and used as models for lupus) turn out spontaneously to make antibodies that react with gp120 even though they have never been exposed to HIV. The British observations (small though the numbers are) are then complementary; an immune response against foreign cells protects against infection with SIV.

Until confirmed (or otherwise), these developments must cause heart-searching throughout the world's programmes of AIDS research. People will be wondering whether the search for a vaccine against HIV should be replaced by programmes to induce tolerance of its protein components. The hunt for mechanisms of the adventitious infections characteristic of AIDS and of Kaposi's sarcoma will be given pause. Even the epidemiologists would be set back on their tracks.

None of this would imply that HIV is irrelevant to AIDS, but that an immune response to foreign cells, most probably lymphocytes, is also necessary. Duesberg will be saying, "I told you so".

John Maddox

Putting a project on ice

Demosthenes Kazanas

ON page 331 of this issue¹ Lowder and colleagues describe the results of a pilot project, carried out in Greenland, to explore the feasibility of constructing an astrophysical neutrino detector in Antarctica. The principles underlying the project are much the same as those behind the established devices of the Kamiokande and Irvine-Michigan-Brookhaven (IMB) experiments, located in mines in Japan and the United States respectively, but with the difference that instead of water it employs a natural ice sheet as the detecting medium.

Neutrino astronomy made its presence felt with a bang a few years ago when the neutrino burst from SN1987A triggered both the Kamiokande and IMB detectors. It has of course been with us for about 20 years now, in a more subdued way, by demonstrating that the measured neutrino flux from the Sun is consistently smaller than that predicted by the standard solar model². Nevertheless, the detection of the neutrinos from SN1987A showed that neutrino astronomy could become a real force in the study of compact objects and gravitational collapse. Unfortunately, the timescale between supernova events in our Galaxy and vicinity is comparable to the lifetimes of most scientists and significant event rates (of the order of one a year) could be achieved only if the detectors became sufficiently large to detect supernova neutrinos from the distance of the Virgo Cluster of galaxies. This would require a detector of approximately 1 megaton (roughly 100 times the size of IMB), a prohibitively expensive project. With stellar-burning neutrino astronomy limited to the study of the Sun, and supernova neutrino astronomy limited in scope by socioeconomic factors, interest has concentrated on the possibility of neutrino astronomy at higher energies of over 1 TeV (10^{12} eV). It is this possibility that Lowder *et al.* address with their feasibility study for AMANDA (Antarctic Muon and Neutrino Detector Array).

The basic principle of devices of this kind is the detection by photomultiplier tubes of Cerenkov radiation created in a large body of water by upward moving muons (in order to discriminate against the background muons, which are produced by cosmic-ray interactions in the atmosphere and move downward). The idea is that the muons are created by interactions of high-energy neutrinos (1 TeV or more) from cosmic sources at large zenith angles (over 90° and preferably close to 180°) either within the detector or in the surrounding water or

rock. Several factors make the detection of higher energy neutrinos economically possible over those of lower energies — the neutrino interaction cross-section increases with energy; the energy available and the range of the resulting muon increases with the energy of the original neutrino, thus requiring fewer photomultipliers to register an event of a given energy; the angle between the neutrino and the muon decreases with energy, thus increasing the possibility of observing point sources above the background; and the background is steeper than the signal expected from the sources, leading to an increase in the signal-to-noise ratio with energy.

The interest in neutrinos of these energies is not new^{3,4}. The possibility was considered in great detail in the 1970s, with a proposal for building a gigaton detector, known as DUMAND⁴, deep underwater off Hawaii. Since then the DUMAND project has been scaled down to roughly 2 megatons, while a number of new proposals have come up for neutrino detectors in the same energy range employing the same technique. The proposals can be divided into those in which the detector would be on the Earth's surface, and those in which it would be underwater. Although it is less easy to work underwater, this disadvantage might be offset by the engineering and higher photomultiplier costs necessary to overcome the higher atmospheric background that is encountered on the surface⁴.

AMANDA is an attempt to combine the best of both worlds in a detector located about 1 km deep in the Antarctic ice. The method of detection is the same with the other detectors of the kind — that is, measurement of Cerenkov radiation by upward moving muons — and Lowder *et al.* show that the absorption length of Cerenkov radiation in polar ice, the crucial property of these detectors, is similar to that of water and thus that ice is an appropriate material to use. The great attraction of AMANDA is that one would be working from a solid surface (rather than in water as in the case of DUMAND), with all the electronic components readily accessible and yet enjoying the advantage of the lower background available at great depths. The necessary photomultipliers can be installed by simply drilling holes in the ice with hot water. Location at the South Pole is also considered an advantage as it would permit observations of sources at constant zenith angle and would allow their simultaneous monitoring by the atmospheric Cerenkov (TeV) and air

shower ($\text{PeV}(10^{15} \text{ eV})$) detectors in the Northern Hemisphere. On the down side, Antarctica is a less congenial and more variable environment in which to work than the waters off Hawaii.

What kind of sources does one expect to observe with such a detector? The answer to this question is at best uncertain, though there are ample theoretical arguments which suggest that there should be a number of astrophysical sources emitting neutrinos at the appropriate energies. The very presence of a diffuse high-energy particle component (the cosmic rays), probably produced in astrophysical shocks⁵, suggests similar particle acceleration in sites where shocks might occur. Possible candidates are accreting compact sources such as X-ray binaries⁶ in our Galaxy or active galactic nuclei which are also presumed to be powered by accretion of matter onto a massive black hole. Lower limits on the expected neutrino fluxes can be estimated from the TeV and PeV fluxes observed from such objects in the past decade⁷, since γ -ray emission at these high energies is expected to be the result of high-energy hadronic interactions and the decay of the copiously produced pions. Unfortunately, further observations in the past few years have cast doubts on the earlier results, indicating that the emission at TeV and PeV energies from accreting binaries is at best sporadic. One should bear in mind, however, that photon-photon interactions within the source can lead to severe attenuation of the high-energy γ -ray emission; thus the absence of γ -ray emission does not necessarily imply the absence of neutrino emission.

The ensemble of active galaxies should also produce a neutrino background easily distinguishable from that of atmospheric events, and individual sources may be distinguishable above this background⁸. The mere detection of neutrinos from these objects would in itself be an important discovery because it would give clues as to the nature of processes that take place deep within their 'central engine', much in the way that supernova neutrinos have allowed us a glimpse, albeit fleeting, of the processes of gravitational collapse of a stellar core. If a significant fraction of the luminosity of these sources is emitted in neutrinos of energy greater than 6.4 PeV, the resonant energy for hadron production via the W particle in antineutrino-electron interactions, more exotic possibilities for detection exist: the combination of the larger available energy with the resonant cross-section and the increase in the effective detector volume to 0.2 km^2 for triggering such cascades leads to a predicted rate of 2,000 events a year with DUMAND from the ensemble of active galaxies⁸.

Also, the possible absorption of neutrinos with energies of over 25 TeV through the Earth's core can yield a value for its density, providing the possibility of the use of neutrinos for geological studies.

Interest in astronomy with TeV neutrino detectors has been gaining momentum over the past decade. Besides DUMAND and AMANDA, there are at present about ten proposals for neutrino detectors of various sizes and at various locations; only last week⁹, the go-ahead was given for the scale-up of Kamiokande, with an expected completion date of 1996. At present, DUMAND seems to be the most advanced of the projects, with an expected deployment of the full 2-megaton array by the end of 1993 at a total cost of about \$10 million.

If positive fluxes are detected, then high-energy neutrino astronomy will become the newest branch of astronomy as we enter the next millennium. □

Demosthenes Kazanas is at the NASA Goddard Space Flight Center, Code 665, Greenbelt, Maryland 20771, USA.

1. Lowder, D. M. *et al. Nature* **353**, 331–333 (1991).
2. Bahcall, J. N. *Neutrino Astrophysics* (Cambridge University Press, 1989).
3. Berenzinsky, V. S. *et al. Astrophysics of Cosmic Rays*, Ch. 8 (North Holland, Amsterdam, 1990).
4. Learned, J. G. *Proc. 15th Texas Symp. Rel. Astrophys., Proc. N. Y. Acad. Sci.* (in the press).
5. Blandford, R. D. & Eichler, D. *Phys. Rep.* **154** 1–75 (1987).
6. Kazanas, D. & Ellison, D. C. *Nature* **319**, 380–382 (1986).
7. Protheroe, R. J. *Proc. 20th ICRC (Moscow)* **8**, 21–44 (1987).
8. Stecker, F. W. *et al. Phys. Rev. Lett.* **66**, 2697–2700 (1991).
9. Swinbanks, D. *Nature* **353**, 200 (1991).

ECOLOGY

Ups and downs in peatland

Peter D. Moore

ONE of the most conspicuous features of the raised mires of the temperate and boreal regions is the stable microtopographic arrangement of hummocks and pools across their surfaces. How is this pattern maintained when conditions such as surface water and peat build-up are continually shifting? An old experimental approach has now been applied to answer this question¹, and in doing so has confounded the ideas of most early peatland ecologists.

A great deal of imaginative thought has been expended on the mechanisms that might account for the constant arrangement of hummocks and hollows scattered or aligned along the contours of a sloping peatland. L. C. Johnson and A. W. H. Damman have now demonstrated that the species of *Sphagnum* bogmosses that occupy the hummocks are more resistant to decay than those that grow in wet hollows and pools¹. The outcome is that the hummocks tend to maintain themselves by more efficient accumulation of peat.

Much of the early work on the vegetation and development of raised bogs was conducted in Scandinavia and it was there that the 'regeneration complex' concept was developed; it was later endorsed by peatland ecologists such as Von Post, Sernander, Osvald, Kulczynski, Tansley and Godwin. Corings and

excavations through peat showed an alternation of pale and dark layers suggestive of an alternation of pool systems with hummocks. The natural supposition was that pools accumulated peat faster because of their waterlogged condition,



Hummocks and hollows — the patterning of mires such as Claish Moss, Argyll, northwest Scotland, is under plant sociological control.

became infilled with time, and eventually developed into hummocks. At this stage the surrounding wetter areas would grow faster and eventually catch up or overtake the current hummocks, so initiating a self-perpetuating, alternating system.

With more detailed peat stratigraphic work, initially in Ireland by Walker and Walker², but subsequently by many others, it became clear that the alternation of hummocks and hollows was not the typical sequence in most raised

mires. More frequent was an approximate maintenance of the positions of hummocks and pools but with occasional increases in pool abundance as a result of changing hydrological conditions on the mire surface, sometimes perhaps as a response to climatic changes. But the question of how the relative stability of pools and hummocks is maintained has remained a contentious issue.

Peat accumulation is essentially the outcome of plant productivity and incomplete microbial decomposition, followed by a degree of compaction as it becomes buried under new layers of peat-forming vegetation. The decomposition part of the equation is the most likely source of variation in peat accretion, and experiments by Clymo³, in which weighed samples of *Sphagnum* moss were placed in bags, buried at different depths in the peat profile and later recovered and reweighed, have shown that the rate of decomposition in a peatland is far greater in the surface layers, where air penetration is frequent, than in deeper layers where more persistent waterlogging and hence anoxia is prevalent. But Clymo also found that hummock-forming species of *Sphagnum*, such as *S. acutifolium*, decayed at similar rates to aquatic species, such as *S. cuspidatum*. However, later experiments involving the use of scanning electron microscopy to trace the degradation of individual leaves⁴, showed that some species, such as *S. magellanicum*, a hummock-former, seemed more resistant to decay than pool species.

Another approach to the question of peatland pattern stability has involved the detailed analysis of hummock and pool peat profiles to determine how effective the process of decomposition has proved at each microsite. Such studies in a raised bog in southern Sweden by Johnson, Damman and Malmer⁵ showed that the *Sphagnum* species in pool peats were more

poorly preserved (measured in terms of intact stem lengths at a given peat depth) than were hummock species at the same depth beneath hummocks. There were also significant differences between the intact survival of stems of *Sphagnum* species within any given microsite. But is this observation the consequence of the relative resistance to decay of the *Sphagnum* species themselves, or is it the outcome of the physical differences between the microsites? The latter possibility seems unlikely as one might expect

poorer decomposition in the waterlogged pool microsites but, on the other hand, local algal photosynthesis could conceivably lead to good oxygenation and therefore rapid decay at the pool water/sediment boundary.

The answer to this question is supplied by Johnson and Damman's experiments at the same site¹. They used the Clymo technique of implanting nylon mesh bags containing samples of *S. cuspidatum* (a pool species) and *S. fuscum* (a hummock species) at a range of depths in both pool and hummock microsites. *S. cuspidatum* was found to lose weight (decay) 1.5–2 times faster than *S. fuscum* both in the hummock locations and in the hollows. The weight loss in both species was influenced by depth in the profile (as Clymo had demonstrated), but was little affected by microsite. Evidently, the different degrees of decay found in the field are a reflection of the relative resistance to decay in the two species.

The implication is that the complex

microtopography found on the surface of actively growing raised peat bogs is maintained by the mosses themselves. Hummock-formers decay poorly, so tend to maintain peat hummocks, while pool species decay more readily and peat accumulates more slowly. The hummock/hollow patterning of such mires is therefore under plant sociological control. The theory of cyclicity, which was born in Sweden, is now firmly laid to rest beneath the peats of that country. □

Peter D. Moore is in the Division of Biosphere Sciences, Kings College London, Campden Hill Road, London W8 7AH, UK.

1. Johnson, L. C. & Damman, A. W. H. *Oikos* **61**, 234–242 (1991).
2. Walker, D. & Walker, P. M. *J. Ecol.* **49**, 169–185 (1961).
3. Clymo, R. S. *J. Ecol.* **53**, 747–758 (1965).
4. Dickinson, C. H. & Maggs, G. H. *New Phytol.* **73**, 1249–1257 (1974).
5. Johnson, L. C., Damman, A. W. H. & Malmer, N. *J. Ecol.* **78**, 633–647 (1990).

ANTIGEN PROCESSING

Proteasomes in the pathway

Miranda Robertson

ATTENTIVE readers of Peter Parham's article in News and Views last December¹ will not be surprised to see the two papers on pages 355 and 357 of this issue^{2,3}. They may also recognize the figure, which shows two putative peptide transporters feeding digested fragments of a cytoplasmic protein into the lumen of the endoplasmic reticulum to form a complex with the two chains of the major histocompatibility complex (MHC) class I molecule. Parham's article was occasioned by the identification in the class II region of the MHC of two genes encoding proteins suspected to be the peptide transporters; but it also contained the prediction that other closely linked genes would shortly prove to encode components of the proteolytic complexes that generate the peptide fragments. The two papers from Brown *et al.*² and Glynn *et al.*³ in this issue, with two others not yet published^{4,5}, bear out that prediction and suggest that the MHC is tantamount to a eukaryotic operon encoding a complete kit for the processing and presentation of antigens to T lymphocytes.

The particular significance of these linked genes in the class II region is for antigen presentation by class I MHC

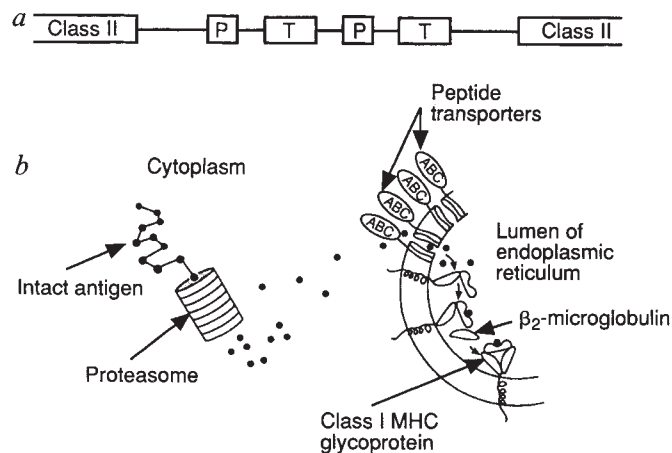
glycoproteins (encoded to the right of the MHC class II region in humans, not shown in the figure), which have to present fragments of proteins produced endogenously by intracellular parasites such as viruses for recognition by cytotoxic T lymphocytes. (Class II glycoproteins specialize in the presentation of antigens synthesized exogenously and in-

ternalized in membrane-bounded vesicles, which presents an equally interesting but quite different problem: see ref 6.) For presentation by class I, therefore, viral proteins synthesized in the cytoplasm must be degraded and transported into the lumen of the endoplasmic reticulum in order to bind the newly synthesized class I heavy chain (see figure). The evidence that two of the non-class-II-glycoprotein genes in the class II region encode peptide transporters is powerful but circumstantial¹. The evidence reported in this issue^{2,3} that two others encode components of a proteolytic enzyme complex is rather better.

These two genes were first identified by Monaco and McDevitt⁷ in mice ten years ago, and in the absence of any clue to their function were designated *LMP-1* and *LMP-2* (for low-molecular-weight proteins). They later turned out to specify the two polymorphic subunits of a large cytoplasmic complex of 16 polypeptides, which Monaco and McDevitt, in the absence of any evidence, speculated as long as five years ago might play a part in antigen processing⁸. Meanwhile, starting from an interest in cellular degradative pathways, Goldberg and colleagues had identified a large proteolytic enzyme complex with a striking ring-like appearance in the electron microscope, and named it the proteasome⁹.

Brown *et al.*² and Ortiz-Navarrete *et al.*⁴ now show that immunoprecipitation of mouse proteasomes followed by two-dimensional gel electrophoresis produces a pattern of spots identical to that given by immunoprecipitation with antibodies against *LMP* products, and that varies with MHC haplotype. Moreover, the genes for *LMP-2*, and for two human genes very closely linked to the transporter genes in the class II region, have been cloned and shown to have sequence similarity to known proteasome components (refs 3 and 5, and J. Trowsdale, personal communication).

The genes encoding the MHC glycoproteins themselves, the transporter genes and the MHC-encoded proteasome genes are all transcriptionally up-regulated by interferon- γ ('immune' interferon), which thus provides a means of coordinately regulating three sets of genes each of whose products is essential to the efficient presentation of antigen. In the absence of antigenic peptides, class I MHC glycoproteins fail to assemble; and mutant cell lines that do not express class I glycopro-



a, Representation of the class II region showing the linkage relationships of proteasome component genes (P) and transporter genes (T), based on information from ref. 3 and J. Trowsdale, personal communication. Other genes in this region are omitted for simplicity. b, An intact antigen is shown binding to a proteasome which degrades the protein in the cytoplasm. The fragments are transported by peptide transporters across the membrane of the endoplasmic reticulum where they bind to newly synthesized class II MHC glycoproteins causing the stable assembly of transmembrane heavy chain with β_2 -microglobulin. Only after the assembly of the three components can the class I molecule be expressed on the cell surface. ABC, ATP-binding cassette transmembrane transporters.

SEMICONDUCTORS

Squeezing light from silicon

Louis Brus

teins, despite having intact genes encoding them, have already been shown to lack at least one of the two transporter genes.

Ortiz-Navarrete *et al.*⁴ have now demonstrated that one of these lines, the human lymphoblastoid line T2, also lacks a proteasome component. The realization that there are two sets of linked genes that may play a part in the delivery of antigens to class I molecules may help to clarify the results of attempts to complement the class I expression defect in cell lines with deletions in the class II region: so far experiments in which mutant cells are transfected with transporter genes have not given consistent results, for reasons that have not been clear. It now seems likely that at least one transporter and one proteasome gene would be necessary for a normal phenotype; and if, as Martinez and Monaco suggest⁵, the MHC-encoded proteasome components serve to bind the proteolytic complex to the genetically linked transporter molecule, it may be necessary to transfect deficient cells with both the transporter gene and the cognate proteasome component gene to restore function. (There is however no evidence yet that proteasomes directly bind to transporters.)

Although only the genes encoding the the MHC glycoproteins themselves are highly polymorphic, there is some polymorphism in both the transporter and the MHC-encoded proteasome genes. These polymorphisms could contribute to the disease associations at present attributed to the MHC glycoproteins. The rat gene *cim*, for example, which maps in the class II region, affects the antigenicity of one of the rat class I molecules in rat strains in which the genes encoding the molecule itself are identical¹⁰. The *cim* gene, the cloning of which has not yet been reported, may therefore be affecting the delivery of peptides to class I molecules. Such interactions between the MHC glycoproteins and the machinery for feeding them with peptides may explain, for example, why disease linkages with MHC glycoproteins so often differ between ethnic groups. □

Miranda Robertson is Biology Editor of Nature.

CRYSTALLINE silicon, the heart of the modern computer and telecommunications industry, is an extremely well characterized material. It has the best combination of physical and chemical properties, making it the material of choice for data processing and memory chips, and it is unlikely to be replaced in this role for a long time. But silicon does not

recombination.

Porous silicon, discovered³ in 1956, is made by electrochemical etching of a silicon anode. For reasons not yet well understood, this etching can create an irregular, three-dimensional microstructure of connected silicon pillars and nodules on length scales down to about 10 nm and below. Whereas bulk silicon

emits quite weakly at its band gap energy, near a wavelength of 1.2 μm in the infrared, this structure emits with higher quantum yield from red to green (Fig. 1).

It has been known for about a decade that the band gap of isolated semiconductor crystallites, of strongly luminescent materials such as CdSe, increases in energy as diameter decreases in the 5 nm range (Fig. 2). This occurs because the electron and hole acquire in small crystallites (quantum crystallites) a significant quantum energy of localization which increases the recombination energy.

In such cases the luminescence is strong in both bulk and quantum crystallite form; silicon is different in that bulk luminescence



FIG. 1 Multicoloured emission from a wafer of porous silicon (L. T. Canham, K. J. Marsh and D. Brumhead).

emit light well, which precludes its use in display devices, lasers for optical communications systems, and so on. Instead, more complex, compound semiconductor materials such as GaAs and InP are used in optical applications.

Therefore one can understand the excitement that followed the discovery by Canham last year that so-called 'porous silicon' emits rather intense visible light following optical excitation¹. The photon energy of this luminescence is much higher (by 1 electronvolt) than the weak emission of bulk silicon. On page 335 of this issue², Cullis and Canham present evidence that quantum confinement is responsible for the new effect.

The luminescence wavelength of a bulk semiconductor is determined by the difference in energy between an excited electron in the conduction band, and a vacancy (hole) left behind in the valence band. This energy is the semiconductor's band gap. When an excited electron recombines with a hole, the energy released appears either as an emitted photon at the band gap energy or as heat. For practical applications, it is important to maximize the luminescence quantum yield, the yield of photons per

recombination. The luminescence should become more allowed in small crystallites as translational symmetry is broken. Luminescence can also increase for purely kinetic reasons related to relaxation dynamics in small crystallites, if the bulk luminescence quantum yield is low. This effect has been observed recently⁴ in the indirect gap semiconductor AgBr.

In fact, last year a group at Cannon Research Center independently discovered⁵ that isolated 3–5-nm Si crystallites in glass emit visible light similar to that from porous silicon, and attributed their observations to quantum confinement. Canham originally proposed that the quantum size effect in porous Si related to the diameter of the silicon wire-like structures. Cullis and Canham now offer strong supporting evidence for this mechanism: first, the silicon structures are primarily crystalline; and second, as the structures become smaller to a diameter of about 3 nm, the emission intensifies and shifts to higher energy. ▶

1. Parham, P. *Nature* **348**, 674–675 (1990).
2. Brown, M. G., Driscoll, J. & Monaco, J. J. *Nature* **353**, 355–357 (1991).
3. Glynn, R. *et al.* *Nature* **353**, 357–360 (1991).
4. Ortiz-Navarrete, V. *et al.* *Nature* (in the press).
5. Martinez, C. K. & Monaco, J. J. *Nature* (in the press).
6. Hackett, C. J. *Nature* **349**, 655–656 (1991).
7. Monaco, J. J. & McDavitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3001–3005 (1981).
8. Monaco, J. J. & McDavitt, H. O. *Hum. Immun.* **15**, 416–426 (1986).
9. Arrigo, A.-P., Tanaka, K., Goldberg, A. L. & Welch, W. J. *Nature* **331**, 192–194 (1988).
10. Livingstone, A. M. *et al.* *J. exp. Med.* **170**, 777–795 (1989).

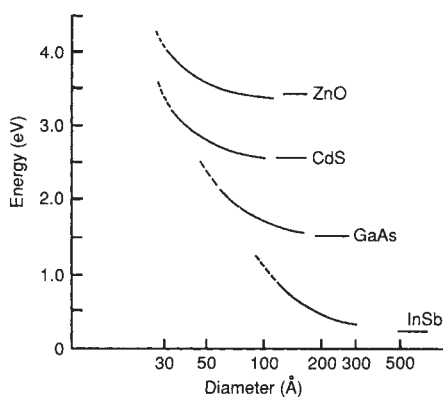


FIG. 2 Calculated dependence⁸ of crystallite band gap (left) on diameter and the bulk gap (right) for several semiconductors.

In all these experiments luminescence is excited optically. While optical characterization is a powerful scientific tool, only electrical excitation would be technologically significant at present. Porous silicon, essentially an irregular Si scaffolding bonded to an Si wafer, is intriguing in this regard. Can the luminescence be excited by electrical current rather than photon absorption? In a remarkable report six months ago, a French group observed a similar visible electroluminescence of a porous Si conducting anode⁶. As etching proceeds, the spectrum intensifies and shifts to higher energy in agreement with the quantum-size hypothesis. In these experiments the holes flow into the scaffolding from the underlying wafer. The electrons are thought to come from oxidation of surface Si-H covalent bonds, creating protons in the electrolyte, and electrons inside the scaffolding. Thus, the surface Si-H bonds serve as sacrificial donors, and the electroluminescence and current flow stop when the surface is completely oxidized.

Many aspects of these phenomena are not understood. The undulating nature of the Si wires suggests that the scaffolding may have both delocalized states that carry current and localized states that possibly emit light. What are the electron-hole dynamics after optical excitation of the scaffolding? Do the electron and hole remain together as a pair, or do they localize separately? Are surface states involved? Can we control the surface chemistry to prevent nonradiative recombination and thus maximize

luminescence? Can the porous Si be intentionally doped to control the conductivity? Can the material be made to show significant optical gain (for laser applications) or stable non-sacrificial electroluminescence?

These experiments also open an unexpected connection between electrochemistry and nanotechnology. It will be useful to understand the etching process

so we may vary the silicon structures. It has been suggested that quantum size effects may themselves control the etching reactions and resulting morphology⁷. After many years on the sidelines, porous silicon may have come of age. □

Louis Brus is at AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA.

RIBOSOMES

Drugs and the RNA world

Harry F. Noller

A NEW strategy of antibiotic action is revealed in a report by von Ahsen, Davies and Schroeder (see page 368 of this issue)¹, who show that the aminoglycoside group of antibiotics, hitherto known as specific inhibitors of translation², can also block self-splicing of group I intron RNA. This discovery of a class of compounds that can inhibit both ribozymes and ribosomes could help to illuminate the primordial past of the modern ribosome.

The discovery of catalytic RNA^{3,4} resolved the long-standing chicken-or-the-egg paradox that troubled early theories of the molecular origins of life: nucleic acids are essential components of any origins scenario, yet they appear to require proteins to function. However if RNA could serve as both gene and functional gene product, then proteins could come later. But the modern translational machinery is so elaborate — requiring more than a hundred distinct macromolecules, among them the large and intricately structured ribosomal RNAs — that it could not have evolved in a small number of steps. On the other hand, it is hard to rationalize gradual, stepwise evolution of even a simplified, primitive ribosome, if its function does not emerge until the final step. More likely, such a structure would have had to evolve from something that pre-existed in the RNA world; the question is, what was it? The work of von Ahsen, Davies and Schroeder¹ could provide us with a clue.

The authors show that a wide range of aminoglycoside antibiotics (which includes the neomycin, kanamycin and gentamycin groups) block the self-splicing reaction of group I intron RNA. The reaction proceeds in two steps, the first of which involves attack by guanosine at the phosphodiester linkage at the 5' end of the intron, followed by a second transesterification reaction in which the 3' hydroxyl of the 5' exon displaces the 3' end of the intron, becoming ligated to the 3' exon in the process⁵. It is the second step of the reaction that is blocked by aminoglyco-

sides, and this occurs at drug concentrations comparable to those that inhibit translation. The relative potency of several antibiotics of the kanamycin group in the self-splicing reaction parallels their ability to inhibit *in vitro* translation.

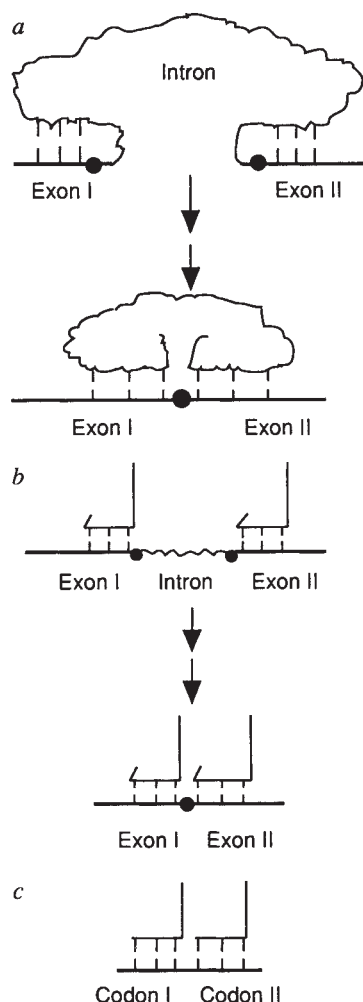
The findings raise the interesting possibility that aminoglycosides may inhibit ribosomes and group I ribozymes by similar mechanisms. Although some of the drugs, such as neamine and paromomycin, are usually potent inhibitors of translation, their effects on ribozyme activity are much weaker, but then some types of ribosomes are resistant to certain members of this group as well.

Mutations in ribosomal proteins as well as in rRNA confer resistance to aminoglycosides². Chemical footprinting has shown that a characteristic cluster of bases is protected in the 1,400–1,500-nucleotide region of 16S rRNA⁶, which is immediately adjacent to the bases protected by aminoacyl transfer RNA⁷. This region of 16S rRNA is also the site of cleavage by the bacterial cytotoxin colicin E3, of a frame-shift suppressor mutation and of several aminoglycoside-resistance mutations, and is most probably the site of codon-anticodon recognition (for review, see ref. 8), which accounts for the increased frequency of translational errors induced by these drugs^{2,9}. It would seem, then, that the decoding site of 16S rRNA is the primary ribosomal target for the aminoglycosides.

This parallel between ribosome and ribozyme could be a coincidence, and the fact that some aminoglycosides bind to multiple sites on ribosomes² shows that this cannot be dismissed out of hand. A more interesting possibility is that certain favoured structural motifs may be common among functional RNAs and that one such motif, shared by 16S rRNA and group I introns, forms a binding site for aminoglycosides. But no primary or secondary structural homology between the two classes of RNA is obvious so far, so perhaps binding specificity is determined by the

1. Canham, L. T. *Appl. Phys. Lett.* **57**, 1046–1048 (1990).
2. Cullis, A. G. & Canham, L. T. *Nature* **353**, 335–338 (1991).
3. Uhlir, A. *Bell Syst. Tech. J.* **35**, 333 (1956).
4. Johansson, K. P., McLendon, G. & Marchetti, A. P. *Chem. Phys. Lett.* **179**, 321–324 (1991).
5. Takagi, H. et al. *Appl. Phys. Lett.* **56**, 2379–2380 (1990).
6. Halimaoui, A. et al. *Appl. Phys. Lett.* **59**, 304–306 (1991).
7. Lehmann, V. & Gosele, U. *Appl. Phys. Lett.* **58**, 856–858 (1991).
8. Brus, L. E. *J. chem. Phys.* **80**, 4403–4409 (1984).

RÉSUMÉ



Base-pairing of short sequences flanking splice junctions in putative ancestral RNAs related to *a*, group I introns or *b*, small nuclear RNAs (snRNAs), could have given rise to complexes that are structurally analogous to two tRNAs bound to adjacent codons on a messenger RNA (*c*). The dots indicate the positions of the splice junctions.

tertiary folding. Davies¹⁰ has made the provocative suggestion that antibiotics may have co-evolved with RNA to produce the modern ribosome, but the idiosyncratic phylogenetic distribution of antibiotic-producing organisms would support more recent origins. Another possibility that comes to mind is that group I introns might have evolved from ribosomes; the fact that they often interrupt rRNA genes hints at such a possibility.

Most intriguing, however, is the converse idea, that ribosomes evolved, at least in part, from an ancestral RNA related to the group I introns. RNA-based life would almost certainly have made use of genetic recombination, which, in the case of RNA, is commonly known as *trans*-splicing. Recognition of splice sites (and, in the case of pre-messenger RNA, branch-point recognition) is usually accomplished by RNA-

RNA base pairing, using short complementary sequences (notable exceptions are pre-tRNA splicing and recognition of the pre-mRNA 3' splice site, which involve proteins). An interesting consequence is that for a generalized primitive splicing reaction, in which the 5' and 3' splice junctions are recognized in this way, the complex at the instant of splicing bears a striking resemblance to two tRNAs bound to mRNA at adjacent codons (see figure). If the transesterification activity of the ribozyme were subverted to catalyse the chemically related acyl-transfer reaction of peptide-bond formation, the stage would be set for the evolution of a primitive ribosome. Selective pressure for such a process could have arisen if the efficiency of splicing were improved by attachment of amino acids or peptides to the ribozyme.

Another link between self-splicing and translation comes from the discovery that the group I introns specifically bind arginine¹¹, which is recognized by the same part of the RNA that binds the guanosine co-factor. Yarus points out that, although the sequence that helps to form the binding site shows considerable phylogenetic variation among the more than one hundred known sequences, it is always centred on a triplet corresponding to one of the arginine codons¹².

Whether or not these findings turn out to be relics of an ancient connection between RNA splicing and the origins of translation remains to be seen. But even if the binding of antibiotics and amino acids to group I introns turns out to be the result of convergent evolution, the implications are no less significant. The evolution of an RNA with the ability to recognize some of the same small molecules that are also recognized by the translation machinery (and to this list can be added the guanosine co-factor) is, after all, not a rare and improbable event. □

Harry F. Noller is at the Sinsheimer Laboratories, University of California, Santa Cruz, California 95064, USA.

Making tracks

HARD on the heels of their molecular shuttle, P. R. Ashton *et al.* have produced a molecular train (*Angew. Chem. Int. Ed. Engl.* **30**, 1042–1045; 1991). Both are supramolecular assemblies based on complexes formed between hydroquinones and bipyridinium cations. One of these species will spontaneously thread itself through a cyclic form of the other, and the resulting complex can be readily converted into two linked macrocycles — a catenane. Ashton *et al.* have now threaded a bipyridinium ring (the train) onto a larger hydroquinone ring (the track) with several docking points (the stations). At room temperature the train runs freely around the track, but at about -60°C it sticks at one station. All that remains is to incorporate into the track signalling devices that can direct the traffic.

On labour day

How does an expectant mother know when to give birth? In the case of sheep at least, the fetus itself signals its readiness for the outside world say T. J. McDonald and P.W. Nathanielsz in the *American Journal of Obstetrics and Gynecology* (**165**, 764–770; 1991). Normally, just before the onset of labour, characteristic changes occur in the levels of three hormones, adrenocorticotrophic hormone and cortisol in the fetus and progesterone in the mother; but these were not seen in cases where a lesion had been made in the paraventricular nucleus of the fetal brain, with the result that pregnancy continued beyond term. How this small region of brain senses and communicates fetal readiness for delivery, and how accurately the sheep model reflects what goes on in humans, are among the remaining questions.

Get knotted

TAKING the biochemical equivalent of doodling to technical heights, J. E. Mueller, S. M. Du and N. C. Seeman report the synthesis of a trefoil knot using single-stranded DNA (*J. Am. chem. Soc.* **113**, 6306–6308; 1991). The authors allow Watson–Crick hydrogen bonding to do the hard work. The single strand they start with contains two distinct pairs of complementary sequences that naturally twine themselves around one another. Once the molecule is suitably entangled, the authors use T4 DNA ligase to join the 3' and 5' ends of the strand to make a single, inextricably knotted loop. They convince themselves that knotted DNA has been made, along with circular and linear molecules, using electrophoresis and sedimentation. As well as producing a three-leafed, singly linked trefoil knot, the authors believe they have made a higher five-leafed, doubly linked form.

1. von Ahsen, U., Davies, J. & Schroder, R. *Nature* **353**, 368–370 (1991).
2. Cundliffe, E. In *The Molecular Basis of Antibiotic Action* (eds Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H. & Waring, M. H.) 402–547 (Wiley, New York, 1981).
3. Cech, T. R., Zaug, A. J. & Grabowski, P. J. *Cell* **27**, 487–496 (1981).
4. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35**, 849–857 (1983).
5. Cech, T. R. *A. Rev. Biochem.* **59**, 543–568 (1990).
6. Moazed, D. & Noller, H. F. *Nature* **327**, 389–394 (1987).
7. Moazed, D. & Noller, H. F. *J. molec. Biol.* **211**, 135–145 (1990).
8. Noller, H. F. *A. Rev. Biochem.* **60**, 191–227 (1991).
9. Davies, J., Gorini, L. & Davis, B. D. *Molec. Pharmacol.* **1**, 93–107 (1965).
10. Davies, J. *Molec. Microbiol.* **4**, 1227–1232 (1990).
11. Yarus, M. *Science* **240**, 1751–1758 (1988).
12. Yarus, M. & Christian, E. L. *Nature* **342**, 349–350 (1990).

Ancient ocean—climate links

Leg 138 shipboard Scientific Party

THE crucial part played by the oceans in regulating global environmental change has led oceanographers over the past 10 years to document past ocean variations on timescales close to the mixing and residence time of nutrients and important biological components of the ocean—climate system. In this way ocean circulation at the end of the last ice-age has been linked to the CO₂ record preserved in polar ice. But only a few select, high-quality sediment cores — recovered by conventional ships — from the top few metres of the seafloor, and so deposited during only the last few glacial cycles, have been studied so far. To gain a more complete understanding of the ocean—climate system we must look back much further, examining its behaviour under many different states. On Leg 138, we set out to apply the high-resolution strategy developed for studies of the past few hundred thousand years to investigate the much older and longer climate record preserved in thick sedimentary sequences that can be recovered only by ocean drilling. We were rewarded with cores spanning 18 million years and recording variations in unprecedented detail.

Leg 138 is a key element of a global study on the role and response of the tropical oceans in climate change on timescales of thousands of years. Five previous legs have sampled the equatorial Atlantic (Leg 108), the Peru Current (Leg 112), the western tropical Indian Ocean (Legs 115 and 117) and the western equatorial Pacific (Leg 130). The eastern equatorial Pacific, which we visited, is critical to this experiment

because it has the greatest potential biological productivity of any of these regions.

Studies of palaeoceanographic records on timescales of a few thousands of years require high sedimentation rates and the recovery of continuous, undisturbed marine sediments spanning millions of years. On previous legs, the common practice has been to drill adjacent holes in an attempt to increase the continuity of recovery. However, even in successive holes drilled with the Advanced Piston Corer (APC), a hydraulic piston corer, significant gaps have been left between core segments. On Leg 138, for the first time real-time laboratory core logging of sediment density (based on γ -ray attenuation), magnetic susceptibility and colour spectral reflectance (with a newly developed instrument) was used to monitor recovered core continuously and to ensure that cores from adjacent holes overlapped at core breaks. By combining real-time monitoring of these high-resolution (2–5 cm) profiles with double- and triple-coring, we were able to construct composite sections and assure nearly complete recovery of the sedimentary section (a single APC-cored hole recovers approximately 90 per cent of the section and a single XCB-cored hole — the extended core barrel, used when the sediment is too stiff for the APC — recovers about 75 per cent of the section).

Using these techniques, we drilled 11 sites, collecting a record-breaking 5,537 m of sediment, along two transects across the eastern equatorial circulation system (Fig. 1). The transects were centred at about 95° W, where the Pacific equatorial circulation is influenced by the eastern boundary of the basin and at 110° W, where the equatorial currents are more zonal and not influenced by the eastern boundary. Sediments spanning the past 7 million years (Myr) were recovered at all sites, with 18-Myr records recovered at three. Although tectonic reconstructions were not a high priority goal of the leg, the basement (or

near-basement) ages determined at 11 sites located on at least three different tectonic plates pointed out several inadequacies with existing plate models and should provide important constraints on reconciling these inadequacies.

The sediment recovered at each site is dominated by the microfossil remains of upper-water-column golden brown algae (nannofossil ooze) with varying amounts

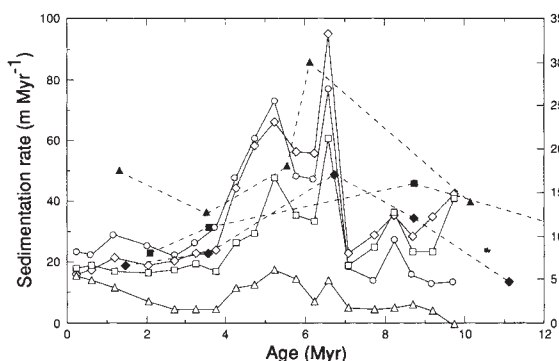


FIG. 2 Generalized sedimentation-rate curves for eastern equatorial Pacific (Leg 138: sites 848, Δ ; 849, $\square$; 850, $\diamond$; 851, $\square$), central equatorial Pacific (Leg 85, site 573, Δ), western Pacific (Leg 130, site 806, $\blacksquare$) and the equatorial Indian Ocean (Leg 115, site 707, $\blacklozenge$). Left-hand scale corresponds to Leg 130 and 138 sites, right-hand scale to Leg 85 and 115 sites.

of phytoplankton and zooplankton, particularly diatom ooze. As is the case today, proximity to the Equator amplified sediment accumulation in the past. Variation in the flux to the seafloor and sediment burial were dominated by temporal changes, however. These often appear to be basin-wide events, typically reflected by changes in calcium carbonate content (probably related to dissolution) or as intervals rich in the diatom *Thalassiothrix longissima*, some of which contain mm-scale laminated beds; the intervals, in millions of years, are <0.65, 1.55–2.1, 3.6–4.5, ~5.5–6.1 and ~8.0–~10.6 (the last two with occasional monospecific intervals).

In the modern ocean, *T. longissima* is generally associated with enhanced upwelling and high productivity and we interpret the presence of these intervals as indicators of major productivity events. Within the limit of our shipboard stratigraphy some of these events appear to be synchronous across the equatorial Pacific and can be imaged with seismic profiling techniques. The broad pattern of temporal variations in sedimentation and accumulation rates is consistent with that found in the central equatorial Pacific (Leg 85), the western equatorial Pacific (Leg 130) and the equatorial Indian Ocean (Leg 115). Comparison of these sedimentation rate curves clearly reveals both the consistency and superior detail of the new curves (Fig. 2).

Superimposed on the long-term changes are higher-frequency fluctuations in the ratio of sedimentary compo-

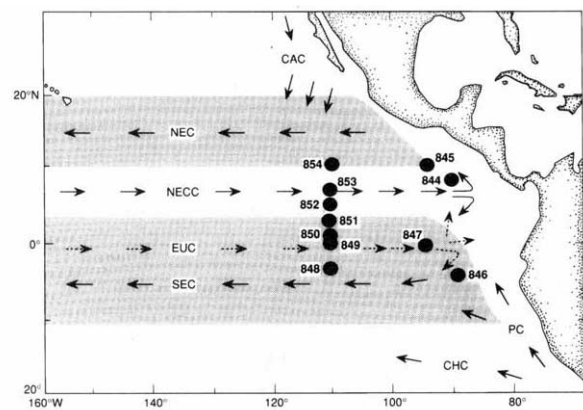


FIG. 1 Leg 138 drill sites and generalized circulation system of the eastern equatorial Pacific. Surface current, solid arrows; subsurface current, dashed arrows. CAC, California Current; NEC, North Equatorial Current; NECC, North Equatorial Countercurrent; EUC, Equatorial Undercurrent; SEC, South Equatorial Current; PC, Peru Current; CHC, Chile Current. Shaded areas illustrate general latitudinal extent of the SEC and NEC.

nents (typically carbonate to silica) which are reflected in the continuous core logs of density, colour and magnetic susceptibility. In particular, the density fluctuations, which for the most part represent changes in carbonate content, show a periodicity that is consistent with Milankovitch climate forcing induced by variations in Earth's orbit. These relationships, in conjunction with the generally excellent bio- and magnetostratigraphy have already allowed us to produce detailed sedimentation and accumulation rate curves and should set the stage for developing a high-resolution stratigraphy. Moreover, the near-continuous density, susceptibility and colour-reflectance logs clearly demonstrate that these detailed records, extending well into the late Neogene, can

be correlated over thousands of kilometres.

It is this framework of detailed and precise correlations that will allow us to construct complete, composite sections for high-resolution sampling. Once analysed, the samples we take, guided by and building on continuous laboratory logs, will provide proxies (such as oxygen and carbon isotopes, carbonate content, grain size, mineralogy and alkenones) for a wide range of ocean-climate system components — ice volume, current and wind patterns, thermal structure and so on. By examining the interrelationships among these proxies we expect to construct the most detailed record of the history of ocean-climate interaction ever attempted for the Pre-Ice-Age world. □

MATERIALS SCIENCE

A new magnetic turn-on

M. B. Salamon

As we tote powerful laptop computers onto aircraft, many of us are keenly aware of the impressive developments in silicon technology. But how many recognize the advances that let us carry ten or twenty times more memory in the form of magnetic disk? Generally regarded as mundane, magnetic recording technology has met the challenges of the micro-computer age, and now boasts of achieving recording densities greater than 100 megabytes per square centimetre. An important component of this technology is a magnetoresistive read-out element in the recording head, usually a thin film that literally flies less than a micrometre above the recording medium. A dramatic improvement in magnetoresistive response, made possible by recent developments by S. S. P. Parkin and colleagues, could herald a further revolution in magnetic recording.

Magnetoresistance, a common property of metals, is the change in electrical resistance that occurs when a magnetic field is applied to the metal. The effect is stronger in ferromagnetic metals, and the material commonly used in magnetic recording is an alloy of iron and nickel. When the magnetization of the metal is reversed from saturation in one direction to its opposite, the resistance change is expressed as a percentage change in total resistance. Changes of 5 per cent are considered excellent for applications. Thus, it should come as no surprise that the discovery of magnetoresistivity changes of 50 per cent and more were greeted with surprise and enthusiasm by those in the business of making magnetic recording devices.

These large magnetoresistive effects were discovered by M. N. Baibich and

coworkers at the University of Paris, Orsay (*Phys. Rev. Lett.* **61**, 2472–2475; 1988). They did not develop an exotic new material, but rather prepared a novel arrangement of ordinary iron and chromium. A superlattice, a sort of club sandwich, was constructed with thin iron layers as the bread and even thinner chromium layer as the filling. When planes of chromium only three atoms thick are used, the magnetizations of successive iron layers, which lie in the plane of the bread, are aligned opposed to one another. A magnetic field of 2 tesla, Baibich and colleagues found, brings the magnetization of the two into alignment and at the same time, reduces the electrical resistance of the film by a factor of two at liquid-helium temperature (4 K).

Unfortunately, the field required to produce this large effect is too high to be of practical use, and the effect is less than a third as large at room temperature. Further, these samples were carefully prepared by molecular-beam epitaxy, a process not amenable to large-scale production. Surprisingly, simple three-layered sandwiches prepared similarly did not show such large effects. Efforts to understand the basis for the large resistance changes identified roughness at the interfaces between the iron and chromium to be essential. Electrons drifting under the influence of the applied electrical field are then forced to pass from one iron layer to the next over a distance so short that they retain the memory of the magnetization direction of the layer from which they last emerged. Electrons pass readily through the interface when the magnetizations of successive layers have parallel align-

ment; increased scattering at these interfaces more than doubles the resistance of the film for antiparallel arrangements.

Now S. S. P. Parkin at IBM and Z. G. Li and D. Smith at Arizona State University have made the needed breakthrough (*Appl. Phys. Lett.* **58**, 2710–2712; 1991). Working with cobalt and copper in place of iron and chromium, respectively, they produced changes in resistance of 65 per cent at room temperature with an alignment field of only 0.2 tesla. A key observation is that the material on which the superlattice is grown, the plate on which the sandwich is served, has a dramatic effect on performance. Films grown on copper were found to be rumpled, with columns of cobalt penetrating through the copper layer. That has the effect of decreasing intended opposite alignment of the cobalt layers. When grown on iron, however, a dramatic improvement in the magnetoresistivity is observed. Parkin *et al.* pursued the effect of underlayers further by using ruthenium. In that case, the magnetoresistive performance was intermediate between that of films grown on iron and copper.

The sample preparation method that Parkin *et al.* use is much more amenable to applications than is molecular beam epitaxy. They begin with a standard silicon wafer, such as might be used for integrated circuit fabrication. Copper, cobalt and iron are deposited on the substrate by means of d.c. magnetron sputtering, a process in which argon atoms are used to erode a source of the desired material, ejecting atoms that then diffuse to the growing film. Computer-controlled shutters are used to prepare a number of identical superlattice structures composed of approximately three atomic layers each of cobalt and copper, 16 pairs in all. The growth temperature is a modest 50 °C. Cobalt, which normally crystallizes in a hexagonal close-packed structure, accommodates itself to the face-centred cubic structure of the iron or copper on which it is deposited. The result is a polycrystalline sample with preferential orientation of three-fold (111) crystalline faces.

The effects observed at room temperature far exceed the magnetoresistance observed in any other material, even the superlattice structure previously studied at cryogenic temperatures. It is fair to expect this achievement to open new possibilities for the use of magnetoresistive materials, not only in magnetic recording, but as sensitive magnetic sensors for other applications. □

M. B. Salamon is in the Materials Research Laboratory, University of Illinois, 104 South Goodwin Avenue, Urbana, Illinois 61801, USA.

A clue to Kartagener's

John Aitken

THE prospect of using transgenic mice to probe the causes of human infertility is raised by Merlino *et al.*, writing in *Genes & Development*¹. The authors have created a transgenic line exhibiting an immotile sperm syndrome that exactly duplicates a flagellar disorder observed in infertile men.

The mice were generated by inserting the complementary DNA for a human epidermal growth factor (EGF) receptor into the mouse genome, driven by the chicken β -actin promoter. One of the resulting transgenic lines (AE24) established a unique pattern of expression, in which extremely high levels of EGF receptor RNA were expressed in the testes but not any other tissue examined. *In situ* hybridization studies indicated that the transgene was transcribed during meiosis, as early as the primary spermatocyte state of spermatogenesis, whereas the appearance of human EGF receptors was delayed until the post-meiotic stages of sperm differentiation. The implication is that the transcription of the transgene was being influenced by a testes-specific enhancer and its expression was being regulated by a translational control mechanism.

The most intriguing aspect of the AE24 line is that about half of the males homozygous for the transgene were infertile. Moreover their infertility was associated with a lack of sperm motility, due to a disorganization of the axoneme (the flagellar central core), frequently involving the loss of microtubule pairs 4-7. This phenotype is identical to a pathology, described by Escalier and David² in cases of human infertility, which is characterized by the presence of spermatozoa in the ejaculate that are normal in every respect except for a loss of motility.

This condition is one of a fascinating number of flagellar mutants in human males that are responsible for infertility. The best characterized is a hereditary condition known as Kartagener's syndrome³, which is associated with the complete absence of dynein arms (the site of ATPase activity) from the axonemes of all ciliary and flagellar structures in the body. As a result, affected patients suffer from chronic sinusitis and bronchiectasis and, more unexpectedly, situs inversus, involving a left-right inversion in the location of key organs such as the heart and liver. Women exhibiting this condition are fertile (so much for ciliary currents wafting eggs down the fallopian tube) but men are sterile, the absence of dynein arms in the sperm tail leading to a complete loss

of sperm motility. Apart from the loss of motility, the spermatozoa from Kartagener's patients are perfectly normal, and if micromanipulated so that they come to lie adjacent to plasma membrane of the oocyte they will fuse with the membrane and fertilize the egg⁴. The possibility of using *in vitro* fertilization technology to treat such patients is therefore being pursued, although whether this is a laudable objective is another question.

Other mutants

In addition to Kartagener's syndrome, a number of other flagellar mutants are known in the infertile male population; they involve the absence of radial spokes, transposition of microtubule doublets from the periphery to the centre of the axoneme (8+2), loss of the central sheath and microtubule deletion. The results of Merlino *et al.* show that a rational approach to understanding the aetiology of these conditions might be to make use of transgenic animals. The important observation that axoneme assembly is normal in the testes is one early fruit of ultrastructural studies of the germ cells generated by the AE24 line; however, microtubule deletion is apparent in spermatozoa located in more distal regions of the reproductive tract, such as the epididymis and vas deferens. In other words, this particular flagellar defect involves a latent instability of the axoneme, rather than a failure to assemble this structure during spermiogenesis. Variations between cells in the rate of microtubule disassembly may explain why the clinical picture is often characterized by variation in the proportion of cells in a given ejaculate expressing a specific flagellar defect⁵.

Transgenic animals may therefore provide an important source of clues about the circumstances responsible for creating the pathological flagellar mutants in man. Such clues are badly needed — the infertile male is the most common cause of intractable human infertility and yet there are few, if any, treatments to tackle this condition. □

John Aitken is in the MRC Reproductive Biology Unit, Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh EH3 9EW, UK.

1. Merlino, G. T. *et al.* *Genes Dev.* **5**, 1395-1406 (1991).
2. Escalier, D. & David, G. *Biol. Cell* **50**, 37-52 (1984).
3. Afzelius, B. A., Eliasson, R., Johnsen, O. & Lindholmer, C. *J. Cell Biol.* **66**, 225-232 (1975).
4. Aitken, R. J., Ross, A. & Lees, M. M. *Fertil. Steril.* **40**, 696-698 (1983).
5. Ryder, T. A., Mobberley, M. A., Hughes L. & Hendry W. F. *Fertil. Steril.* **53**, 556-560 (1990).

Skimmed water

THE hovercraft is a good idea poorly implemented. It floats elegantly on a cushion of compressed air; but it compresses that air, and propels and steers itself, with clumsy airscrews. Daedalus now proposes improvements.

Firstly, he says, the air cushion should be compressed not with a fan, but directly from the craft's forward motion. The envelope of his new 'ramcraft' skims the water all round it, except in front, where it rises into an air-intake. Air scooped in by forward motion is compressed and trapped by the ram effect, and supports the craft. With careful aerodynamic design, almost the full theoretical stagnation-pressure should be reached under the envelope (compared to some 20 per cent of it under a typical aircraft wing). Even so, at low speeds the resulting lift will only be a few kilograms per square metre. This modest 'wing-loading' enforces a very light style of construction, and Daedalus's design is a sort of cross between a tent and a hang-glider. Its big fabric wing gapes high in front to form the air-inlet, but curves down to the ground at the back and sides to trap the air. A few light guy-ropes and spars define its shape. Stationary, the ramcraft will be a sad and droopy sight; but when it starts to move, the ram-pressure of entering air will soon inflate it to shape, and lift it off. At a few metres a second, it should be fully inflated and airborne. Daringly, the passenger and control cabin hangs beneath the wing itself, in the pressurized region. Passengers will not even notice the small over-pressure.

The ramcraft will propel itself over water in flying-fish style. This creature often skims along with its tail still submerged, and wagging fast to maintain speed. So the ramcraft will simply extend a propeller into the water on a long angled shaft, and steer and propel itself from the thrust. No noisy, inefficient airscrews are needed.

To let the ramcraft ride safely up onto a beach or apron, Daedalus will fit the blade-tips of its propeller with small angled traction-rollers. They will then roll smoothly over solid ground, continuing to propel the craft by a form of linear worm-drive action.

Daedalus will initially launch the ramcraft as a small personal fun-vehicle, combining the scary delights of hang-gliding and water-skiing. But like balloons and other tension structures, and unlike conventional boats and aircraft, it should get better as it gets bigger. Huge bulk-carrier ramcraft, little more than acres of brightly-coloured, self-inflated fabric, may yet challenge the supertanker for dominion of the seas. DAVID JONES

Under the volcano

SIR — In making a case for the inadequacies of the Japanese government's volcano prediction committee, David Swinbanks (*Nature* 351, 511; 1991) suggests that the eruption of Mount Unzen, because it was not predicted in advance, resulted in the death of 42 people. But the eruption had been under way for two weeks, the tragic deaths being caused by the avalanching of hot debris from the front of a thick lava dome, perhaps triggered by a small explosion near the summit of the volcano.

This dome had been slowly advancing over the headwall of a steep canyon on the east side of the volcano, and the resulting avalanches had been front-page news in Japan for more than a week before the disaster. Network television channels (I was in Japan at the time) carried spectacular images of hot lava blocks spalling from the oversteepened dome and cartwheeling down the canyon to form pyroclastic flows. The volcanologists and Japanese TV crews who perished knew full well that pyroclastic flows were being triggered. Why else would they have been up there taking pictures? I do not agree that the statements of the Japanese prediction committee were connected in any way with their demise. In confusing the difference between predicting eruptions (which is sometimes possible) and predicting the generation of hot avalanches (which is virtually impossible), Swinbanks unfairly blames the Japanese interagency committee because it did not "fully anticipate the particular threat that claimed [42] lives". I believe that no committee could have predicted the magnitude and timing of the 3 June avalanche and resulting deadly pyroclastic flow.

Swinbanks acknowledges that the volcano eruption committee, chaired by Professor Daisuke Shimozuru, did in fact warn that pyroclastic flows were apt to occur. But he chides the committee for not explaining "what a pyroclastic flow is, how deadly it is, or how fast it can strike". Such a statement is simply unfair. Are we to criticize a flood prediction committee for not explicitly warning that rising waters might cause people to drown? People in Japan are well aware of the dangers of pyroclastic flows, thanks to excellent programmes in public understanding of science, as well as the numerous TV newscasts at the time. People did not need an interagency committee to tell them these avalanches are dangerous or that they move fast.

Because the early parts of the eruption took place during Japan's rainy season, swiftly moving mudflows were (and continue to be) a significant hazard at Unzen. No committee or group should be

criticized for emphasizing the destructive potential of volcanic mudflows. On 26 May, when it became obvious that parts of the growing lava dome had already become unstable, the Unzen volcano prediction committee stated that pyroclastic flows might also occur. Thousands of people were safely evacuated in the next few days; the only people to perish on 3 June were the 42 individuals, including 15 journalists and 3 volcanologists, who elected to take the risk of entering the restricted zone. It is worth remembering that this zone was established in large part on the basis of statements from the Unzen volcano prediction committee. This is a sad and paradoxical story, but not one of inadequate performance by that committee.

Two final points: the inset map accompanying an article by Christopher Anderson on the same page as Swinbanks' mistakenly locates Sakura-jima volcano on the Shimabara peninsula, next to Unzen. Sakura-jima is actually about 140 km to the south. And the correct spelling of the US Geological Survey researcher cited is Tom Casadevall.

RICHARD S. FISKE
*National Museum of Natural History,
Smithsonian Institution,
Washington, DC 20560, USA*

SIR — Many volcanologists around the world are dismayed about the disparaging tone of David Swinbanks' News article about volcano prediction by Japanese scientists. Japanese volcanologists are among the world leaders in developing predictive techniques for dealing with threatening volcanoes. In keeping with their well-established tradition, they closely monitored the activity of Unzen volcano throughout the current volcanic crisis, which began in November 1990, and they have accurately assessed its eruptive potential based on analyses of both the history of the volcano and the characteristics of its new unrest. Their fine performance stands in stark contrast to the astonishing implications in Swinbanks' account.

Despite warnings by Japanese scientists and successful evacuation of vulnerable areas, 42 lives were lost at Unzen in early June in a pyroclastic flow that was considerably larger than those preceding it. Seeking a scapegoat for the tragedy, some journalists, including Swinbanks, identified and blamed the volcano prediction committee. But a more thorough investigation would have revealed that after long dormancy of a volcano (200 years at Unzen), the eventual outcome of renewed unrest is notoriously difficult to predict. Nevertheless, Japanese researchers had correctly assessed the nature of

the immediate hazards at Unzen. Further, Japan has developed one of the world's most effective systems of collaboration among the civil government, scientists and law-enforcement and emergency-management agencies for dealing with natural hazards. Their procedures are models for other nations, and these procedures were operating effectively at Unzen.

The people who died so tragically had knowingly entered the clearly designated danger zone, generally for professional reasons: among the victims were three experienced volcanologists who fully understood the hazards that they faced. The deaths cannot and should not be ascribed to any deficiency in the performance of the Japanese volcano advisory committee. Swinbanks expressed surprise that the Japanese public continues to listen to warnings: if, however, local people had failed to heed the warnings at Unzen, without doubt the casualty toll would have been much higher.

DONALD W. PETERSON
(IAVCEI)
250 Davenport Way,
Palo Alto,
California 94306, USA

SWINBANKS REPLIES — I did not say that the volcano prediction committee failed to predict the pyroclastic flow of 3 June. Rather, I argued that the committee failed fully to anticipate the danger of pyroclastic flows and failed to issue adequate public warnings.

The first pyroclastic flows occurred at Unzen on 24 May. But despite the known dangers of such flows, the committee did not comment on them until after a major flow on 26 May. The committee's statement said only: "the occurrence of pyroclastic and mud flows can be expected from now on and a strict alert is necessary".

Fiske argues that pyroclastic flows are a well-known phenomenon in Japan, and thus it was not necessary for the committee to explain their dangers. But pyroclastic flows have not occurred at Unzen for at least 200 years, nor are they common elsewhere in Japan.

After the 3 June tragedy, pyroclastic flow (*kasairyu*) became a household word in Japan. But in the critical period of late May media coverage of the phenomenon was much more limited. Surely one of the roles of the committee should be to educate the public?

The committee has perhaps opened itself to some undue criticism by calling itself a "prediction" committee. As Professor Shimozuru explains on page 295, volcanic prediction is a bit like "betting". Given the limitations of the science, should not the committee be renamed so as not to raise false expectations in the general public? □

Images at the blind spot

SIR — Ramachandran and Gregory¹ explored the intriguing question of how we "fill in"² that part of a scene which is not sampled by the visual system — one example is the blind spot, discovered by Mariotte in 1660. Helmholtz³ reviewed the early literature but was unable to put forward a satisfactory theory.

We have devised a new method to investigate the response to images on and near the blind spot. Using a desktop computer, contiguous black squares each subtending 20' at the eye were successively displayed on the screen, forming a horizontal line which gradually extended from one side of the screen to the other at 2° per second. A fixation point was provided.

Initially, one sees clearly the line starting at one side of the screen and lengthening towards the centre (see figure). When it arrives at the edge of the blind spot, all movement ceases and the line appears to maintain its position and length; despite the fact that squares continue to be added and their images cover more and more of the blind spot. Eventually, when a square exits the other edge of the blind spot, it is seen as a continuation of the original line which had stopped. The line is suddenly seen to lengthen once more. The duration for

which the movement is stopped equals the time required for the squares to move a distance equal to the diameter of the blind spot, as measured by conventional means (5°).

In the second experiment, two stimuli were used, each identical to those in the first experiment but with the lines separated vertically so that the image of one passed above (or below) the optic nerve head. The appearances of each of the two resulting lines were dramatically different. The line that traversed the blind spot was seen to be shorter. For example, if the total subtense of the lines were 10°, the line traversing the blind spot was seen to be shorter by 5°, half the length of the comparison line that missed the blind spot. These effects are so pronounced that they are readily and instantly perceived by the most inexperienced of subjects, presumably because they involve relative movement as opposed to more conventional, static demonstrations. Our technique may be diagnostically valuable, particularly as the movement would make the patient's task of reporting much easier. A copy of the program discussed above is available from us.

P. R. ANDREWS
F. W. CAMPBELL

*Physiological Laboratory,
University of Cambridge,
Downing St, Cambridge CB2 3EG, UK*

1. Ramachandran, V. S. & Gregory, R. L. *Nature* **350**, 699–702 (1991).
2. Rudolf, N. D. de M. & Wilson, M. E. *King's College Hospital Gazette* **37**, 1 (1958).
3. Helmholtz, H. *Treatise on Physiological Optics* Vol II (ed. Southall, J. P. C.) 170–220 (Dover, New York, 1962).

Muller's ratchet and flu virus

SIR — According to Chao¹, the fitness of RNA viruses can be decreased by Muller's ratchet. Chao's findings in a bacteriophage system are also applicable to influenza virus populations.

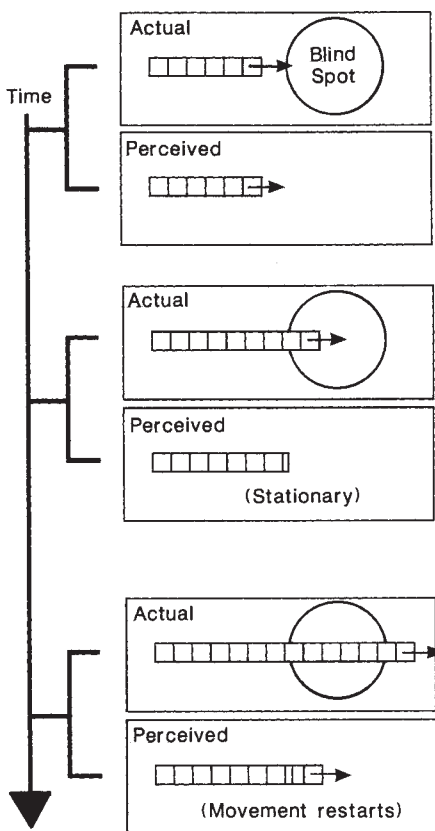
The ratchet operates in viral populations which undergo periodic bottlenecks. Following infection with influenza virus, for example, individuals develop a humoral immune response which protects them from reinfection with the same virus. After an influenza virus strain has circulated extensively, most of the world's human population becomes immune to reinfection². Thus the number of susceptible individuals to support influenza virus replication is drastically reduced and with it the pool of human influenza virus genes: the immune response imposes the bottleneck. Eventually, influenza virus genomes containing critical mutations on the genes for the surface glycoproteins which allow them to evade the immune response, arise. These viruses, usually one or a few

particles, are able to reinfect previously immune individuals and act as propagules. At this point, the human influenza virus population escapes the bottleneck and the progeny of one or a few genomes will become the prevalent virus population, as found by Chao.

Nucleotide sequence and phylogenetic analyses of each of the eight RNA segments of the genome of flu virus isolated from humans and animals during the past 40 years have led to the construction of evolutionary trees of the influenza virus genes^{3,4}. Detailed analyses of the evolution of these genes reveal that the selective pressure of the immune response on haemagglutinin (HA) and neuraminidase (NA) genes results in the fixation of random mutations in these genes and in other viral genome segments, such as segment 8 (encoding nonstructural proteins NS1 and NS2) or segment 1 (encoding basic polymerase PB2)^{3,4}. The genes encoding PB2, NS and nucleoprotein (NP) of human influenza viruses have accumulated many silent and nonsilent mutations in their genomes as a consequence of population bottlenecks imposed by immune selection. Decreased fitness of the derived (recent) influenza genomes would be predicted if Muller's ratchet applied to this human influenza virus–host system.

The re-emergence of an ancestral fossil genome ('Russian' strains, similar in nucleotide sequence to A/Fort Warren/50, dating from 1950) in the human population in 1977 offers the possibility of testing this prediction. After the reappearance of this fossil there was ample opportunity for reassortment of genome segments. Viruses such as A/California/78 and A/Kiev/59/79, which have genomes consisting of a mixture of genes from the circulating viruses and the reintroduced fossil virus, were isolated from the human population during two consecutive seasons. Were Muller's ratchet applicable, the fossil influenza virus genome segments reintroduced in 1977 would have displaced by reassortment the more derived and less fit circulating gene segments (evolved from the 1950 virus ancestor). Recent isolates of influenza virus from humans do indeed have the derived NS, NP and PB2 genes with cumulative mutations and not the fossil genes, suggesting equal or greater fitness in the derived genes. Thus the reintroduced virus is still in circulation but it did not displace the previously circulating virus lineages.

These findings can be reconciled with those of Chao in the bacteriophage sys-



Actual and perceived movement at three successive instants.

1. Chao, L. *Nature* **348**, 454 (1990).
2. Couch, R. B. & Kasel, J. A. *Rev. Microbiol.* **37**, 529–549 (1983).
3. Buonagurio, D. A., Nakada, S., Parvin, J. D., Krystal, M., Palese, P. & Fitch, W. M. *Science* **232**, 980–982 (1986).
4. Gorman, O. T., Donis, R. O., Kawaoka, Y. & Webster, R. G. *J. Virol.* **64**, 4893–4902 (1990).

tem by assuming that bacteriophage $\phi 6$ has evolved to reach the fittest genome to replicate in its host whereas human influenza virus has not. Because the fittest flu genome has yet not evolved, the accumulation of mutations as a consequence of population bottlenecks yields a more fit genome. In other words, the ratchet operates only on genomes that have reached maximum fitness.

RUBEN O. DONIS

Department of Veterinary Science,
University of Nebraska,
Lincoln, Nebraska 68583, USA

Sexual semantics

SIR — J. L. Martinez (*Nature* 352, 288; 1991) objects to the use of the term 'sex' to describe the diverse recombination processes in bacteria (J. Maynard Smith *et al.* *Nature* 349, 29–31; 1991), preferring to reserve the term for "the well-characterized mechanism of reproduction-linked genetic recombination in eukaryotes". But if bacteria don't have sex, neither do many eukaryotes.

Many eukaryotes are protists, many of whom reproduce without sex, have sex without reproducing, and have sex with themselves. To assume that recombination in bacteria is "completely unrelated to what we call sex" is to ignore the fact that protists (and bacteria) tell us that reproduction, reduction division, crossing-over, ploidy and syngamy are separate phenomena, which may or may not have evolved under similar selective forces. This is precisely why these groups are so interesting, and why we should adopt a liberal attitude towards 'sex'.

INNES C. CUTHILL

Department of Zoology,
University of Bristol,
Bristol BS8 1UG, UK

SIR — Martinez¹ doubts whether it is appropriate to use the term 'sex' in the context of bacteria and would restrict its use to "the well-characterized mechanism of reproduction-linked genetic recombination in eukaryotes, leaving bacteria with their own very different mechanisms of genetic recombination which are usually completely unrelated to what we usually call sex".

Precisely this problem beset us when compiling entries for a biology dictionary². We hope we were helpful when we wrote the entry for sexual reproduction (capitals indicate cross-references), which begins as follows: "In its broadest sense, any process by which genetic material is transferred from one cell, and/or individual, to another (see CONJUGATION); or any process involving genetic RECOMBINATION. Such a view of sexual reproduction would have to include PLASMID trans-

fer, viral infection, antibody production and even some forms of GENE MANIPULATION. One possible benefit of throwing the concept wide open this way is that it may reveal possible clues to the origin of SEX. The processes of *sex* (recombination and/or genetic transfer) and *reproduction* (production of new individuals) are sometimes only tenuously linked (e.g. in PARTHENOGENESIS in *Paramecium*), and may even be uncoupled . . ."

In the entry on SEX we point out that this term is often used as a synonym of sexual reproduction, and attempt to distinguish its several other senses — most of which involve organisms where different sexes or sex organs are identifiable.

We welcome suggestions for improvements to these and other entries.

MICHAEL THAIN

Harrow School,
Harrow,
Middlesex HA1 3HW, UK

1. Martinez, J. L. *Nature* 352, 288 (1991).

2. Abercrombie, M., Hickman, M., Johnson, M. L. & Thain, M. *The New Penguin Dictionary of Biology* (Penguin, London 1990).

Ocean uncertainty

SIR — Numerical climate models typically predict that doubling atmospheric CO₂ would lead to eventual, or equilibrium, global warming of 2–5 K. How quickly the warming would occur depends in part on the behaviour of the oceans, which influence the rate of greenhouse warming by taking up both heat and CO₂ from the atmosphere. Considering the heat-uptake question alone, studies with coupled oceanic-atmospheric general circulation models make it clear that the rate at which the climate would change is model-dependent to an even greater extent than is the equilibrium warming. Further, a recent simulation suggests that much of the uncertainty in rate of climatic change may be due to uncertainty in oceanic rather than atmospheric processes.

The table shows lag times derived from published results of coupled oceanic-atmospheric general circulation models, each of which was subjected to a sudden step-function doubling of atmospheric CO₂. Although the lag time of the real climate system depends on the history of greenhouse gas concentrations, the idealized case of sudden CO₂ doubling provides an upper limit to the spread of model response times under a common scheme of greenhouse forcing. In the table, ΔT_{eq} denotes the equilibrium surface air warming due to doubled CO₂, and τ_e denotes the *e*-folding relaxation time — the time it takes after CO₂ is doubled to reach

MODEL RESPONSES TO DOUBLED CO₂

Model	Ref.	ΔT_{eq} (K)	τ_e (yr)	κ_{eff} (cm ² s ⁻¹)
GISS	5	4.2	124	2.6
OSU	6	2.8	46	2.1
NCAR	7	3.5?	49	1.5?
GFDL	8	2.4	42	2.7
OPYC	9	3.0	14	0.6
+ECHAM				

Question mark, uncertainty about whether equilibrium was reached.

$1 - e^{-1} \approx 63\%$ of ΔT_{eq} . Among this set of models, ΔT_{eq} varies by less than a factor of 2, but τ_e varies by nearly an order of magnitude.

The right-hand column of the table is an attempt to identify the part of the uncertainty in lag time due to the oceanic component of the coupled models. As shown by Hansen *et al.*¹ and Wigley and Schlesinger², if downward heat flux in the oceans is proportional to the vertical temperature gradient, so that the ocean in effect transports heat by diffusion, then τ_e scales as $(\Delta T_{eq})^2$. To separate out this ΔT_{eq} -dependence, the right-hand column gives the effective thermal diffusivity κ_{eff} of the oceans according to Wigley and Schlesinger's analytical model²: $\tau_e \approx (2.74 \text{ cm}^2 \text{ s}^{-1} \text{ K}^{-1}) (\kappa_{eff}) (\Delta T_{eq})^2$. If the analytical ocean model is modified to account for heat transport due to the sinking of cold water at the poles (thermohaline circulation), the dependence of τ_e on ΔT_{eq} is less steep than second power³, so the range of uncertainty in oceanic processes may be underestimated by the range of κ_{eff} among the models.

I find that κ_{eff} varies by less than a factor of two among the GISS, NCAR, OSU and GFDL simulations, suggesting that most of the variation in τ_e among these models is actually due to variations in ΔT_{eq} . The GISS simulation uses a simplified, purely diffusive ocean model whereas the remaining three use Cartesian-coordinate oceanic general circulation models of the type developed two decades ago at GFDL. But the final entry, a simulation with the Hamburg version of the ECMWF atmospheric model coupled with a new 'isopycnic' oceanic model developed by J. M. Oberhuber, obtains a significantly faster ocean response than the other models even if it is assumed that τ_e scales as $(\Delta T_{eq})^2$. Although this particular result is highly preliminary, the quick surface warming may be attributable to the careful restriction of heat diffusion across surfaces of constant density in Oberhuber's model, which alone of all the ocean models shown in the table uses density as a vertical coordinate.

Narrowing uncertainty in the rate of climatic change is important for compar-

ing theories of greenhouse warming with Earth's temperature record over the past century and for estimating the response to time-evolving concentrations of greenhouse gases next century. I believe the results I have discussed make a case for diagnosis and intercomparison of oceanic general circulation models, analogous to comparison of atmospheric models⁴, with emphasis on rate of response to external perturbations.

CURT COVEY

*Atmospheric and Geophysical Sciences
Division,
Lawrence Livermore National
Laboratory,
Livermore, California 94551, USA*

1. Hansen, J. *et al.* *Science* **229**, 857–859 (1985).
2. Wigley, T. M. L. & Schlesinger, M. E. *Nature* **315**, 649–652 (1985).
3. Harvey, L. D. D. *J. geophys. Res.* **91**, 2709–2717 (1986).
4. Cess, R. D. *et al.* *Science* **245**, 513–516 (1989).
5. Hansen, J. *et al.* in *Climate Processes and Climate Sensitivity* (ed Hansen, J. & Takahashi, T.) 130–162 (American Geophysical Union, Washington, 1984).
6. Schlesinger, M. E. *Climate Dynamics* **1**, 35–51 (1986).
7. Washington, W. M. & Meehl, G. A. *Climate Dynamics* **4**, 1–38 (1989).
8. Manabe, S., Bryan, K. & Spelman, M. J. *J. phys. Oceanogr.* **20**, 722–749 (1990).
9. Sausen, R., Lunkeit, F. & Oberhuber, J. M. in *Research Activities in Atmospheric and Oceanic Modeling* (ed. Boer, G. J.) 9.4–9.5 (WMO, Geneva, 1990).

Rhythms of war

SIR — Schreiber *et al.* proposed¹ that seasonal rhythms influence the opening dates of wars, based on the observed significant correlation between daily photoperiod and the dates of commencement of acts of aggression². They assigned this correlation to increased aggression induced by long photoperiods. Their published correlations were indeed quite remarkable for data from both the Northern and Southern Hemispheres and from the Equator.

However, as with all such comparisons and correlations, an understanding of underlying principles and of the nature of the dataset should be applied to the interpretation of the results. What this high correlation may in fact identify is a relationship between the season and military strategy 'over the past few thousand years'. It seems logical that military strategists should make use of good conditions (including long photoperiod) to move men and equipment, and that they would plan campaigns accordingly. The failure to respond to climate conditions while planning and executing wars has been illustrated by more than one disastrous campaign.

ROBERT BENNETT

*Imperial Cancer Research Fund,
London WC2A 3PX, UK*

1. Schreiber, G., Avissar, S. & Tzahor, Z. *Nature* **349**, 574–575 (1991).
2. Laffin, J. *Brassey's Battles: 3500 Years of Conflict* (Pergamon, London, 1986).

Krill abundance

SIR — Antarctic krill (*Euphausia superba*) is the primary food source for many animals in the southern ocean and is also the basis of a large fishery. To manage this resource, krill abundance has been estimated directly with acoustics and indirectly with estimates of predator consumption rates. However, the abundance estimates using acoustics are often an order of magnitude less than those based on predator consumption rates¹. The acoustic method converts echo energy to absolute biomass by assuming that the echo return is the sum of individual scatterers, and by assuming an empirical or modelled acoustic target strength (TS) for individual krill. Everson *et al.*² and Greene *et al.*³ reported new TS measurements of experimentally constrained krill, but to date no corroborating field data have been published. We present new *in situ* TS measurements of krill obtained in March 1991 off Elephant Island, Antarctica.

Foote *et al.*¹ ensonified live krill aggregations in a cage at 120 kHz. The mean single-animal target strength of krill (lengths 30–39 mm) was inferred from the aggregation backscatter to range from –81 to –74 dB. Past acoustic surveys have typically used TS values

calculated using the equations from the BIOMASS⁴ program, but these equations lead to gross underestimates of krill abundance².

Until recently, a fluid sphere model was thought to characterize adequately the acoustic TS of krill. Wiebe *et al.*⁵ ensonified several species of live, but tethered, zooplankton at 420 KHz and concluded that sound scatter from elongated animals is better described by a bent cylinder model⁶ and that TS is proportional to the volume of an animal rather than its cross-sectional area. Using these data, Greene *et al.*³ predicted krill TS at several frequencies and over a range of animal lengths. The Foote *et al.* data agree with the TS prediction for a mean animal size of 33 mm. Because the bent cylinder model predicts that TS is much more sensitive to animal length than would be expected with the fluid sphere model, additional measurements of krill TS, particularly at different lengths, could provide strong corroborations of refs 1 and 3.

We present *in-situ* TS measurements of krill (mean length = 47.44 mm, $\sigma = 2.92$), around Elephant Island (*a* in the figure). Zooplankton sound scattering depends upon the morphology of the animal as well as its size, shape and orientation⁵. The spread of the distribution is likely due to the size distribution and variable orientation of the ensonified krill. Some of the high TS observations may be from multiple krill erroneously identified as individual scatterers. Nonetheless, the distribution is centred on –69 dB, within 1 dB of the prediction by Greene *et al.* Furthermore, the slope between the Foote *et al.* data and the data presented here is in accordance with that predicted by the bent cylinder model.

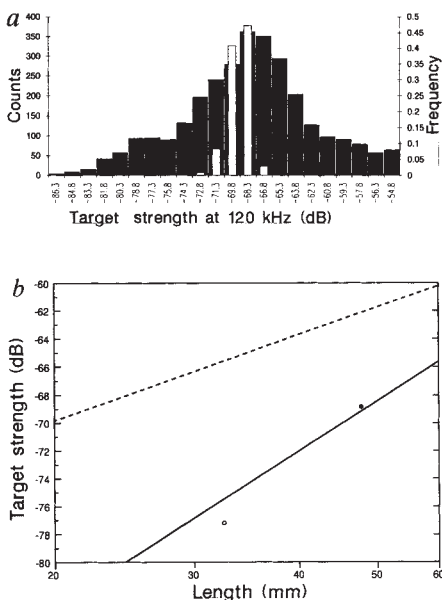
The TS of 47.44-mm krill estimated from the BIOMASS equations is –62.3 dB, 6.7 dB above the modal value of the measurements reported here. Thus, for a population composed of 47.44-mm individuals, abundance estimates using the mean TS reported here would be 4.7 times higher than that estimated with the BIOMASS value.

ROGER P. HEWITT

*Southwest Fisheries Science Center,
La Jolla, California 92038, USA*

DAVID A. DEMER

*Scripps Institution of Oceanography,
University of California at San Diego
La Jolla, California 92038, USA*



a, TS distribution for 2,957 individual krill detected with a Simrad EK500 echo sounder and a 120-kHz split-beam transducer off Elephant Island, Antarctica in March 1991. Solid bars, *in situ* measurements; open bars, distribution predicted from the sampled distribution of animal lengths. *b*, TS by length relationship at 120 kHz. Solid line reproduced from ref. 3. Also plotted are the median length and TS from the Foote *et al.* experiment (open circle)¹, the data reported here (solid circle), and the BIOMASS equation (dotted line).

1. Foote, K. G., Everson, I., Watkins, J. L. & Bone, D. G. *Acoust. Soc. Am.* **87**, 16–24 (1990).
2. Everson, I., Watkins, J. L., Bone, D. G. & Foote, K. G. *Nature* **345**, 338–340 (1990).
3. Greene, C. H., Wiebe, P. H. & McClatchie, S. *Nature*, **349**, 110 (1991).
4. Post-Fibex Acoustic Workshop *BIOMASS Report Ser. No.* 40 (1986).
5. Wiebe, P. H., Greene, C. H., Stanton, T. K. & Burczynski, J. *J. Acoust. Soc. Am.* **88**, 2346–2360 (1990).
6. Stanton, T. K. *J. Acoust. Soc. Am.* **86**, 1499 (1989).

Structural revolution

M. F. Perutz

Introduction to Protein Structure. By Carl Branden and John Tooze. *Garland: 1991.* Pp. 302. \$27.95 (pbk).

In the mid-1940s, when I had to teach students the atomic structure of minerals, I marvelled that in a mere 20 years X-ray crystallography had ordered the bewildering variety of external forms into limited sets of structural patterns. Fifteen years later, when the structures of myoglobin and haemoglobin were found to be similar, I hoped that in the fullness of time X-ray crystallography would do proteins the same service. Branden and Tooze's lucidly written and beautifully illustrated volume shows that my expectations have now been at least partially fulfilled.

α -Helices, β -strands and β -turns are woven into recurrent structural motifs that may be either purely helical, as in haemerythrin or the globins, or pure β , as in the immunoglobulins, or mixed, as in the barrel of the enzyme triosephosphate isomerase. Some motifs, such as the one that binds calcium ions, can be recognized by their sequence alone; others, like the β -barrel of the virus coat proteins, share almost no sequence similarity.

A third of the book is devoted to proteins that bind nucleic acids or nucleotides. The two chapters on transcription factors illustrate the convergent evolution of similar structural motifs for similar functions in strikingly different species: a wonderful example is the analogy between the DNA-binding folds of the phage λ repressor and the homeodomain of *Drosophila*. It is typical of the rapid progress in this field, rather than a criticism of the book, that the hypothetical structures of two DNA-protein complexes described in the book have already been overtaken, and in one instance disproved, by high-resolution X-ray studies.

Immunoglobulins and the class I molecules of the major histocompatibility complex are discussed together with detailed illustrations of their interactions with haptens and antigens. Most of the chapter on membrane proteins is rightly devoted to the structure and function of the photochemical reaction centre of

Rhodospseudomonas. I was surprised to find only 15 pages devoted to enzyme catalysis, with serine proteases as the sole examples, and no reference to the activation of their zymogens or their inhibition by serpins. The discussion on prediction, engineering and design of proteins is well balanced and is highlighted by Kaspar Kirschner's amazingly accurate prediction of the structure of the α -subunit of tryptophan synthetase from its sequence alone. Bond energies that stabilize protein structures are analysed on the basis of Brian Matthews' directed mutagenesis of T4 phage lysozyme, and bond energies between enzymes and substrates with the help of Alan

sages will be invaluable for students.

I have some minor criticisms. The first chapter omits the fourth important type of secondary structure, the polyproline helix which is the structural basis of all bones and connective tissue. The only allosteric protein referred to is the Trp repressor, a poor example, because its two subunits change their tertiary structure independently on binding tryptophan, without the cooperativity and change in quaternary structure that are the hallmarks of allostery. Moreover, the authors omit the intriguing stereochemical mechanism of the structural change, as though this were trivial or too hard for students to understand. This vital omission brings me to my main criticism: the authors seem to have been so carried away by the attractions of their topological diagrams that they have often omitted the chemistry that gives them meaning. They must also have heeded the warning given to Stephen Hawking when writing *A Brief History of Time* that each equation

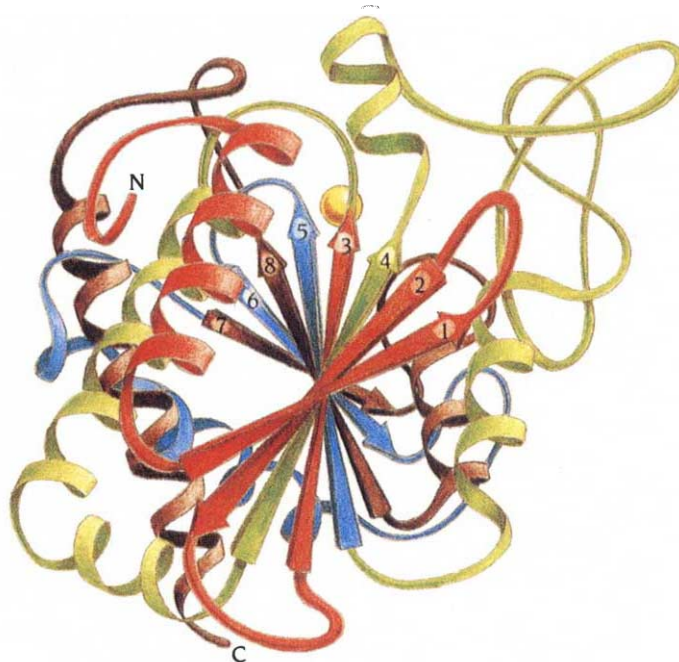
would halve the book's sales. Their book contains neither mathematical nor chemical equations. The section on X-ray analysis glosses over the fundamental concept of Fourier transforms.

Headings of chapters and sections invariably take the categorical form of 'all cats have whiskers'. I do not like this J. D. Watsonesque style, because it gives students the false impression that everything in science is cut and dried and has only to be memorized, like law. For example, the authors state that heterozygotes for sickle-cell anaemia are resistant to malaria because the malarial parasite lowers the pH of the host erythrocyte sufficiently to cause sickling, resulting in a loss of potassium ions that kills the parasite. But this explanation cannot be correct

because the erythrocytes contain only about 40 per cent sickle-cell haemoglobin S. They do not sickle, even if the pH drops, and therefore do not lose potassium ions.

Despite these blemishes, Branden and Tooze's book provides a most attractive and readable introduction to the many new molecular forms in the living world that are being unravelled at an ever increasing pace by X rays and NMR. $\square$

M. F. Perutz is in the Laboratory of Molecular Biology, Medical Research Council, Hills Road, Cambridge CB2 2QH, UK.



Pretty protein: diagram of the structure of carboxypeptidase, which consists of a mixed β -sheet of eight β -strands with α -helices on both sides, and binds a catalytically essential zinc atom (yellow circle).

Fersht's studies of the complex between tyrosyl-tRNA synthetase and tyrosyl-AMP. The book ends with a 15-page introduction to structure determination by X-ray crystallography and NMR.

The strength of the book lies in its beautiful art work and its logical dissection of the baffling complexities of protein structures. In addition, it contains several admirably concise, lucid and accurate presentations of difficult concepts, such as the genetic basis of antibody diversity, electron transport by the photochemical reaction centre or structure determination by NMR. These pas-

Models of mind

John Fox

Unified Theories of Cognition. By Allen Newell. *Harvard University Press: 1990.* Pp. 549. \$39.95, £31.95.

ARTIFICIAL intelligence is often said to date from a meeting at Dartmouth College in 1956. Among the most remembered contributions to the meeting was a computer program called the Logic Theorist. Allen Newell, one of its authors, has remained a prominent figure in artificial intelligence ever since, influencing much of its theory, practice and philosophy. He has also been consistently fascinated with natural intelligence. *Unified Theories of Cognition* is based on Newell's 1987 William James Lectures and expresses these preoccupations. It is a statement of the weight to be expected from a scientist of his calibre and experience, but it is not simply the retrospective that many, late in a productive career, might settle for. The book ranges authoritatively over 40 years of research in psychology, artificial intelligence and computer science, but in the service of a new and challenging thesis — that cognitive science is on the threshold of a unifying theory of mind.

Or more precisely, it can be if it wants to be.

A rush of interest in artificial intelligence in the early 1980s led to an influx into the field of people with little appreciation of natural science — computer technologists, mathematicians and logicians, and countless practical programmers. The subject probably gained in formal rigour but it also lost some of its relevance and excitement for many scientists more interested in the natural phenomena of mind.

Newell wants to put some of the excitement back. To this end he adopts a double strategy. First, he argues that psychology, particularly experimental psychology, has yielded enough solid facts that emphasis can now be given to pulling them together theoretically. Second, he tries to show that this ambitious, even controversial, aim is realistic by setting out a particular style of theorizing which he exemplifies with 'SOAR', a theory of the human cognitive architecture.

Newell begins by arguing that unified theories of cognition are desirable and possible. He then presents his highly individual views on computers and computation, and on intelligence and the processes that underpin it. The next chapter ranges widely from neurophysiology to deliberative problem solving (and beyond). Here Newell establishes the terms in which empirical results are to come into contact with computational concepts. He focuses on the timescales of mental phenomena, distinguishing a dozen or so 'bands' — the main ones discussed are the 'biological', 'cognitive' and 'rational'. These might be thought of as merely levels of description, but Newell holds that they reflect discontinuities that must emerge given the kind of computational system we believe the brain to be. He identifies these bands as the loci of distinct types of empirical phenomena.

Newell then moves on to present his candidate theory. SOAR is an architecture, a fixed information-processing structure specialized to raise and achieve goals adaptively, by exploiting a body of knowledge, and acquiring new knowledge in the process. SOAR operates in cycles, repeatedly retrieving knowledge relevant to its current goal and making decisions on the results. Decisions are made about what knowledge to use in achieving its goals ('problem spaces'), what interpretations to assign to data ('states') and what 'operators' to apply. If SOAR has suitable knowledge, the goal will be achieved. If not, it will 'impasse'. Impasses cause SOAR to generate new, impasse-specific goals automatically, and to decide recursively on problem spaces, states, operators and so on. If the subgoal is achieved, the

result can be used to reach the parent goal. After resolution of an impasse, the architecture identifies the critical contextual data and 'chunks' this as a new fragment of knowledge for future use.

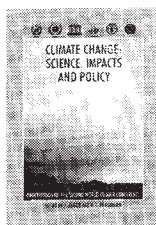
In the next three chapters, Newell considers various topics for which there is a good stock of reliable empirical results, discussing how SOAR can explain them. Among these are choice reaction times, skilled typing, learning and skill acquisition, and logical reasoning. Where necessary, he is candid about any problems the data raise for the theory. The final chapter is equally open about larger topics (such as natural language and development), which the theory does not yet attempt to address. For theorists who might hope to do better, he concludes with some recommendations for how to proceed.

The book is occasionally hard-going and sometimes inconsistent in its assumptions about the reader's knowledge; but in general it is lively, lucid and persuasive. Notwithstanding Newell's own caveats, the explanatory power of this superficially simple theory seems impressive. But other reviewers will be better qualified to assess the empirical adequacy of SOAR. I want instead to focus on Newell's central belief that SOAR exemplifies a substantial and important advance in the methodology of cognitive science, and make some comments on how the SOAR research programme is developing in practice.

Newell is clear about explanations and methods in psychology. He is lakatosian rather than popperian: there are no critical experiments, only progressive accumulation of results that strengthen or weaken belief in specific theories. Unified theories are huge undertakings so we must explicitly support scientific communities to pursue them. But community means communication and hence shared languages, perhaps technical ones. Here one might think that the mathematicians and logicians mentioned earlier could contribute their notations and techniques, but Newell seems sceptical about their contribution: "Only when strong simplifications are made . . . is mathematics potent enough to be useful".

A substantial international community has begun to participate in the development and empirical testing of SOAR. Versions of SOAR are distributed to members as programs (large ones) accompanied by weighty documentation, a profusion of research papers and so on. (The book describes version four; at the present time a greatly revised version five is standard.) The SOAR community works incredibly hard at attracting, training and supporting recruits. But the community has little interest in formal presentation or discussion, and it seems

Climate Change: Science, Impacts and Policy



Proceedings of the Second World Climate Conference

Edited by J. JÄGER and H. L. FERGUSON

This book is not only an important statement about the current knowledge of climate change and its effect on the world's environment and prospects for sustainable socio-economic development, but also a summary of the technical programmes and political actions needed to deal with this major global problem.

£50.00 net/\$95.00 HB 0 521 41631 0 592pp. 1991
£24.95 net/\$44.95 PB 0 521 42630 8



The Edinburgh Building,
Cambridge CB2 2RU, UK.

40 West 20th Street, New York,
NY 10011-4211, USA.

to me that a trick is being missed here; all this effort is needed precisely because the community has not adopted unambiguous, public languages and notations for expressing its ideas. I cannot agree with Newell that mathematical notations necessitate oversimplification (though I do not imagine that developing appropriate formal languages will be easy), and it seems to me that the methodology runs a number of risks.

Some of my colleagues say that they find SOAR's details impenetrable. This book helps a lot, but SOAR community members have been heard to remark that to understand the theory properly, a considerable investment in using the program is needed. There must be a danger here that assessment by uncommitted scientists is discouraged. I even wonder if active SOAR workers, who presumably have a good grip on the theory, are constrained from participating in its evolution. SOAR software is developed centrally; satellites can influence the centre but it must be difficult for them to explore theory revisions independently by experimenting with modifications to the program.

There is much in SOAR to interest the wider artificial-intelligence community, but the lack of formality and unfamiliarity of language can produce puzzlement, even scepticism. I take just one example, from Newell's own list of important topics that SOAR does not yet address — decision-making under risk and uncertainty. SOAR makes decisions, its states and operators bear some relationship to 'beliefs' and 'actions', for which there are some well-understood logics and calculi. But their precise nature is unclear, so any insights from other frameworks are hard to exploit. On the other hand, the powerful impasse idea is unknown to logicians and decision theorists. The two communities have much to offer each other, but a wall is up.

The back cover of *Unified Theories of Cognition* includes much fulsome praise from the eminent. Before I read the book I thought this was over the top, but now I am not so sure. I am convinced, indeed excited, by Newell's vision, and impressed by the SOAR community's achievements. But I am not convinced that its practical methodology is yet strong enough for other, perhaps less able, scientists to develop and compare alternative theories. SOAR is either an important success or an important failure, but I think we shall only know which, and why, when we have better ways of expressing and discussing such ambitious proposals. □

John Fox is in the Advanced Computation Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

Homo habilis laid bare

Bernard Wood

Olduvai Gorge. Vol. 4. The Skulls, Endocasts and Teeth of *Homo habilis*. (Two books.) By P. V. Tobias. Cambridge University Press: 1991. Pp. 921. £110, \$175.

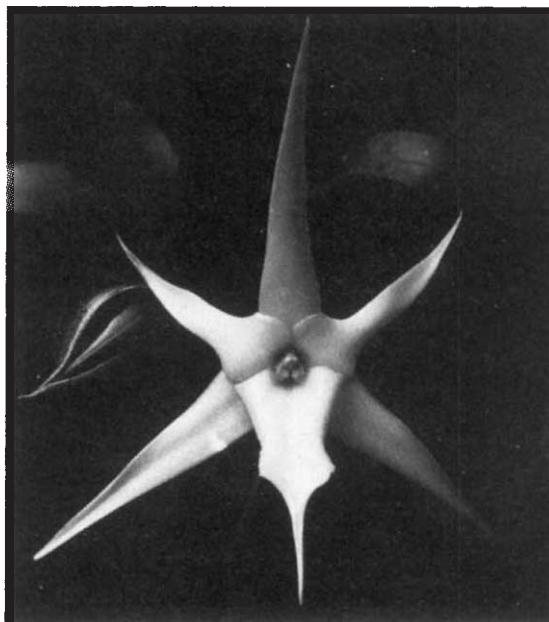
OLDUVAI Gorge in Tanzania is pre-eminent among early hominid sites. Other locations have yielded more and better-preserved hominid fossils, but 'the gorge' holds a unique place in the history of palaeoanthropology. Its importance rests both on the nature of the hominid fossils found there, and on the profound influence that work at Olduvai has had on the general development of palaeoanthropological field studies and laboratory research. It has proved to be a testing ground for methods of investigation and analysis that have subsequently been adopted for the exploration of other important early hominid sites at Omo and Koobi Fora in the Turkana Basin, North Kenya, and at Hadar, further north in the Gregory Rift.

Before the announcement of the dis-

covery of a 'robust' australopithecine cranium, OH5, at Olduvai in 1959, virtually all of the important evidence for early human evolution was confined to discoveries made at sites in the south of the African continent. These cave sites, at Taung, Sterkfontein, Makapansgat and Swartkrans, had furnished copious evidence of small-brained australopithecines, or 'man-apes' as they became known, as well as more fragmentary evidence of a hominid not unlike *Homo erectus*, a creature best-known from sites in Indonesia and China. The '1959' Olduvai cranium, initially attributed to *Zinjanthropus boisei* but now usually referred to as *Australopithecus boisei*, demonstrated that australopithecines were also found in East Africa, although the substantial morphological gap between the australopithecines and *Homo erectus* remained unfilled. This void was not to last for long, for the extra research supported by funds made available following the discovery of OH5 soon yielded results. The discovery in 1959 of a single tooth, OH4, alerted Louis Leakey to the possibility of the presence of a second hominid at Olduvai; the discovery in 1960 of OH7 and, in 1963, of OH13, confirmed his suspicions.

Thus began the realization that Olduvai had provided evidence of not just

Flowers on film



graph of a plant in the text-books is as rare as that of a human face is common". Not until the turn of the century did the extensive photographic plant survey take a hold, and not until the 1920s and 30s, with the works of Karl Blossfeldt and Ernst Fuhrmann, did it become commonplace. For this reason, the photography of Scowen & Co in Ceylon during the 1870s to 90s, an example of which is shown here, is exceptional. The 'orchidomania' that was sweeping through Europe in the 1870s probably motivated Charles Scowen to record for tourists the exotic varieties he found in the botanical gardens at Peradeniya, although the pictorial and technical excellence of the photographs suggest that Europe's horticulturalists were also the intended buyers. This photograph, taken around 1870, is of *Angraecum sesquipedale* and comes from the magnificent collection of images reproduced in *Flora Photographica: Masterpieces of Flower Photography from 1835 to the Present* by William A. Ewing. Published by Thames & Hudson at £24.95.

P.T.

one new hominid taxon, *Australopithecus boisei*, but two, the second being *Homo habilis*. Louis and Mary Leakey had already retained the anatomical advice and services of Professor Phillip Tobias to help them interpret OH5, and to describe and analyse it in detail. When the possible importance of the subsequent discoveries was appreciated, Tobias's brief was extended to studying the skulls and teeth of what was to constitute the latest new taxon. The Leakeys' faith in the scientific talents of 'PVT', as he is widely known, was amply justified. His monograph on the OH5 cranium was a seminal work in palaeoanthropology. It set new standards for the analysis of morphological detail, and many of the phylogenetic interpretations it espoused proved to be durable and influential. Now, nearly a quarter of a century later, the same incomparable anatomical talents have been turned towards the task of describing and analysing the items that make up the collection of *Homo habilis* cranial remains from Olduvai.

It is surely appropriate, and not hyperbole, to use the term *magnum opus* to refer to this work. The sole authorship by Tobias of more than 900 pages, arranged in two tomes and weighing 5.2 kilograms, is a formidable achievement. His text begins with a summary of the history of ideas about *Homo habilis*. He chronicles the initial scepticism of colleagues and then, with justifiable satisfaction, goes on to record recent evidence that the hominine status of *Homo habilis* is now widely, if not universally, accepted. Important specimens are dealt with in order of completeness, beginning with the cranium OH24, then turning to OH13, OH16 and OH7. These cranial remains, together with the mandible OH37, and several sets of isolated teeth, dental and calvarial fragments, comprise the 19 specimens that Tobias considers in the volume.

The level of morphological detail corresponds to that in Tobias's previous volume in this series, but in this case the structure of the work, which deals with an apparently homogeneous collection specimen-by-specimen instead of region-by-region, does make the monograph unwieldy, both literally and metaphorically. For example, the reader needs to concentrate hard to locate the whereabouts of reviews of the remains attributed to each tooth type. Tables of measurements occur throughout the text, placing an additional, and probably unnecessary, strain on both the patience and the wrists of the reader.

We have come to expect a high standard of accuracy from the author, but the fact that he achieves, and maintains, that standard throughout is remarkable nonetheless. The upper limit of the cranial

capacity of *Homo habilis* is incorrectly quoted; but in general, factual infelicities are commendably rare. The unwary may be confused by the term 'gens'. It reads as if it may have been the result of a well-intentioned, but disastrous, spell-check procedure for genus, but Tobias explains that he uses this little-used label to denote a phyletic lineage. Aside from the detailed analysis of the preserved morphology, the text includes useful contextual information about both site localities and the location of residual distortion in specimens such as OH24. Tobias's warning that the basi-bregmatic and supraorbital heights of this specimen are underestimates, and Clarke's caution, set out in the second of the appendices, that its cranial capacity in the undistorted state "must inevitably be greater than the absolute capacity as measured now" (page 858), should be heeded by those workers who seek to interpret the affinities of this important

specimen.

As fine a work of scholarship as this is, it is unlikely to be the last word on *Homo habilis*. Comparisons with relevant remains found elsewhere in Africa are relatively few and developments since 1985 are largely unreported. Tobias's analytical methods are traditional, and his approach is statistically relatively unsophisticated. But his meticulous documentation of the morphological evidence is sufficient reason for this latest monograph to become the definitive and enduring record of the important fossil hominids from Olduvai Gorge. The case for Tobias to be numbered among the masters of palaeoanthropology was already a strong one, but with the publication of these volumes, his place in their midst is assured. □

Bernard Wood is in the Department of Human Anatomy and Cell Biology, University of Liverpool, Liverpool L69 3BX, UK.

Bridging the quantum gap

L. Eaves

Quantum Semiconductor Structures: Fundamentals and Applications. By Claude Weisbuch and Borge Vinter. Academic: 1991. Pp.252. \$34.95, £25.

THE physics of quantum semiconductor structures is now one of the most lively areas of condensed-matter research. The quantum Hall effect, where the Hall resistivity changes in steps of h/e^2 , is perhaps the best-known property of such devices, as its discovery led to Klaus von Klitzing's Nobel prize and to a new international standard of electrical resistance. In addition, the devices made from quantum semiconductor structures have entered many homes in the form of semiconductor lasers to read the music on compact discs and fast transistors at the front end of instruments to receive satellite broadcasting. Although much of the underlying physics is fairly straightforward, there has been a notable and frustrating absence of any text providing a general but thorough introduction to the field. This excellent new book admirably fills the gap, no doubt in part due to the fact that its authors work at one of the leading European laboratories for research and development of semiconductor devices, Thompson CSF in Paris.

The quantum structures described fall into two main categories. In the first, the electrons are constrained to move in a plane ('two-dimensional electrons'). By depositing the dopant atoms well away from this plane, the motion of the elec-

trons is unimpeded by scattering from these impurities. This results in long electronic mean-free paths when the structures are cooled to liquid-helium temperatures. Under these conditions, and especially when a magnetic field is applied, the electrical resistance of the structures is dominated by the quantum-mechanical behaviour of the conduction electrons.

In the second type of device, the electronic states and optical properties are governed by the electronic motion and quantum tunnelling in the direction perpendicular to the layered planes of a semiconductor heterostructure — essentially a multiple sandwich structure composed of layers of different semiconducting materials. The authors give a clear and detailed account of quantum wells, resonant tunnelling devices and superlattices, and of how these can be used to make new types of electronic and optoelectronic devices such as quantum-well lasers, optical modulators and very-high-frequency devices.

The text is well illustrated by many figures and diagrams. It will be essential reading for postgraduate research students entering this field and also for those studying university courses in physics and electronic engineering. Experienced researchers will also find many of the explanations illuminating and the bibliography an excellent source of references. □

L. Eaves is in the Department of Physics, University of Nottingham, Nottingham NG7 2RD, UK.

New Journals

Appearing in next week's issue is *Nature's* New Journals supplement, in which more than 50 journals are to be reviewed.

The origin of the cosmic X-ray background

T. Shanks*, I. Georgantopoulos†, G. C. Stewart†, K. A. Pounds†,
B. J. Boyle‡ & R. E. Griffiths§

* Physics Department, University of Durham, South Road, Durham DH1 3LE, UK

† Department of Physics & Astronomy, University of Leicester, Leicester LE1 7RH, UK

‡ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

§ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore Maryland 21218, USA

A high-resolution image from the Rosat X-ray satellite reveals many faint discrete sources in the 0.1–2-keV energy range. Optical spectroscopy of these sources performed at the Anglo-Australian Telescope shows that many of them are quasars, and the inferred density of quasars on the sky contributes at least 30% of the cosmic X-ray background at 1 keV.

The full field of the Rosat image (diameter 2°), integrated over the entire PSPC energy band, is shown in Fig. 1. Within the central area of diameter $40'$, the detector resolution is reasonably uniform at $25''$ (full width at half maximum), and sensitivity is relatively constant. At larger angles from the axis, the image becomes more blurred, and sensitivity falls; we therefore confine our analyses to the central area where the signal-to-noise ratio for source detections is highest. Lists of sources were obtained using both the UK and the German analysis software which include both 'sliding box' and point-spread-function fitting procedures. These algorithms were applied to the data in soft (<0.5 keV), hard (>0.5 keV) and total energy bands. Within the aperture of radius $30''$ used for the 'sliding box', the average counts expected from background were 16, 4 and 20 in the soft, hard and total energy bands respectively. The sources have a detection significance corresponding to $>4\sigma$ in at least one band with a minimum of 10 counts. With these detection criteria we estimate that, statistically, less than one spurious source should arise in the central region of $40'$ diameter. For a power-law spectrum with an energy index of 1 the conversion factor from counts in the total band to flux in the energy range $S(0.1\text{--}2.4\text{ keV})$ is $1.0 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$ per (Rosat count s^{-1}). This value is sensitive to the exact spectral form of the sources, and so we

THE extragalactic X-ray background was discovered more than 20 years ago, but its origin still remains unclear. Some authors^{1–3} have suggested that up to 100% of the extragalactic background could be due to active galactic nuclei (AGN) whereas others⁴ have suggested upper limits as low as 30%. Theoretical predictions (for example refs 5, 6) which estimated the X-ray luminosity function from the well determined optical luminosity function suggested that the likely contribution from quasi-stellar objects (QSOs) was $\sim 50\%$. This estimate is based, however, on the uncertain X-ray properties of faint QSOs. Other possible contributors to the background that have been suggested over the years include hot intergalactic gas, X-ray-emitting galaxy clusters and starburst galaxies (for example, see refs 7, 8 and 9).

Previous attempts to resolve the discrete source component directly include the work of Griffiths *et al.*¹⁰ who identified sources in the Pavo field of the Einstein Deep Survey. In four Einstein high-resolution imager exposures, each lasting 90,000 s, and subsequent optical follow-up, they discovered 14 QSOs, corresponding to a surface density of $32 \pm 9 \text{ deg}^{-2}$. Other Einstein surveys were less deep¹¹ and thus found lower surface densities. Our first aim is therefore to determine whether the high QSO surface density found in the Pavo field is typical of other high-latitude fields. In fact, in our deeper observation we find an X-ray QSO surface density $\sim 2\text{--}3$ times as high as that found in the Pavo survey.

Rosat observations

A long, 30,000-s exposure was made with the position-sensitive proportional counter (PSPC) on Rosat, with the satellite in pointed mode, in late July 1990, before the start of the Rosat all-sky survey. The PSPC detector^{12,13} is sensitive in the 0.1–2.4-keV band and has similar light gathering capability to the Einstein imaging proportional counter (IPC). But it has considerably lower instrumental background and superior spatial resolution which lead to greater sensitivity, together with much improved spectral resolution.

The field selected for observation was the QSF3 field (03 h 41 min, -44°) of Boyle *et al.*¹⁴. We have systematically surveyed this field for QSOs using the optical ultraviolet excess (UVX) technique, following this up with optical fibre spectroscopy. These observations identified 19 UVX QSOs with redshift $z < 2.2$ as well as three compact, narrow-emission-line galaxies in a field of $40'$ diameter. The average surface density¹⁵ of QSOs of magnitude brighter than $B = 21.2$ (the lower limit for detection in the QSF3 field survey) is 41 deg^{-2} , and so the observed density is within 1.3σ of the mean.

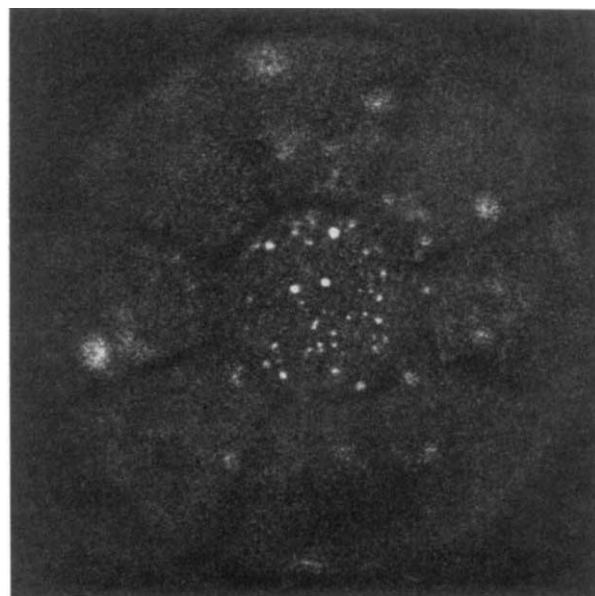


FIG. 1 The full field of the Rosat PSPC image, 2° in diameter. The energy band is 0.1–2.4 keV and the faintest source detected has a flux of $\sim 9 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$. The dark annulus (radius $20'$) and lines are due to obstruction of the field by the window support structure. Thirty-nine sources have been detected in the central area of diameter $40'$, leading to a sky density of sources of 112 deg^{-2} . The X-ray brightest source is a $B < 11$ mag star, and the two next brightest are UVX QSOs with $z < 1$.

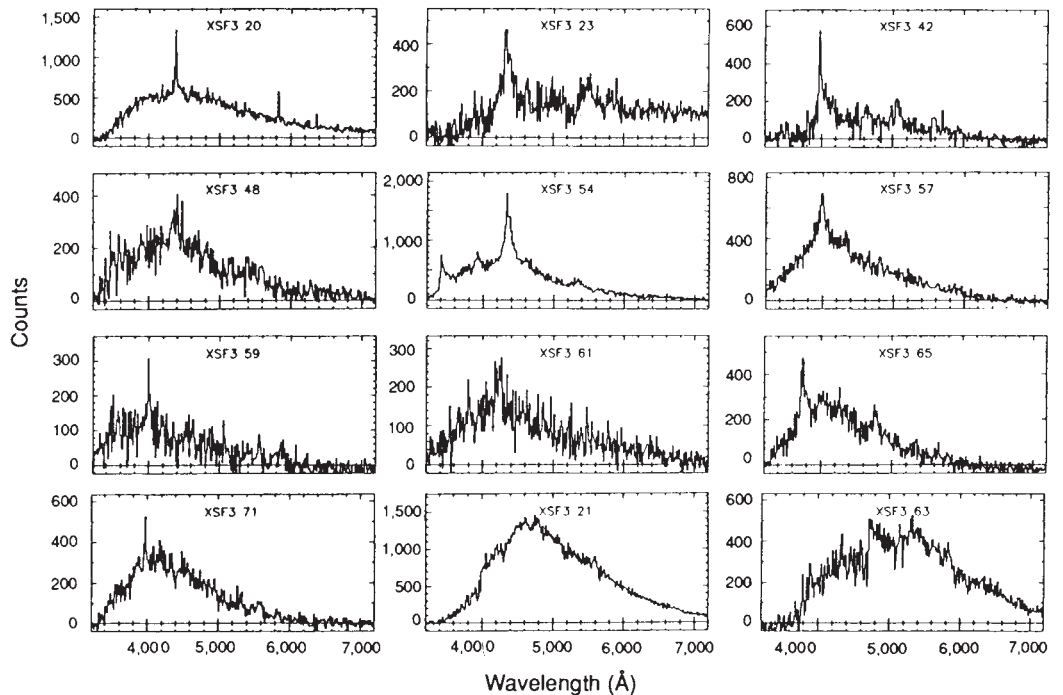


FIG. 2 The AAT Autofib optical spectra of 10 newly confirmed QSO X-ray sources in the QSF3 field. Also shown in the bottom panel is one of the galactic stars (XSF3:21) and the $z=0.18$ E galaxy (XSF3:63) which were also detected by Rosat.

shall confine our later detailed analysis to the 0.5–2-keV band where the appropriate conversion factor is 1.2×10^{-11} in the same units.

In the above bands, the $\sim 4\sigma$ limits for source detection at the edge of the central region are $S(0.1\text{--}2.4\text{ keV}) \approx 2 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ and $S(0.5\text{--}2\text{ keV}) \approx 1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$; the significance levels at these flux limits are a little higher for sources nearer the field centre because of the slightly improved resolution. The source positions, fluxes at 0.1–2.4 keV and detection significances are listed in Table 1.

A total of 39 sources were detected in the central area of 0.35 deg^2 , corresponding to an overall source surface density of $112 \pm 18 \text{ deg}^{-2}$. The faintest source in the survey has a flux of $S(0.1\text{--}2.4\text{ keV}) = 8.8 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$. Two further possible sources may not have been detected at the formal 4σ limit because they lie close to other sources; one lies $\sim 1'$ north of XSF3:57, the other $\sim 1'$ south of XSF3:65. Although these may have higher fluxes than some listed sources, to be conservative we have not included these in any of our analyses. Despite some sources visually appearing extended, all the sources in Table 1 are statistically consistent with being pointlike. The relative positional accuracy of sources is $\sim \pm 6''$. The Rosat source density should be compared with 8 deg^{-2} in the Extended Medium-Sensitivity Survey (EMSS)¹⁶ and 36 deg^{-2} in the Pavo field of the Einstein Deep Survey¹⁰.

Optical observations

We used the COSMOS machine measurements of the IIIaJ UK Schmidt plate (J2672) previously used for the UVX survey to search for optical counterparts of the 39 X-ray sources. After correcting for small, systematic offsets in right ascension and declination in the absolute Rosat positions, we identified 12 of the X-ray sources with QSOs from the UVX survey to an accuracy of $\pm 5''$. Of the seven UVX QSOs that were undetected by Rosat, one (QSF3:18) lies close to the obstructed annulus on the PSPC, and another (QSF3:12) is only $45''$ from the brightest X-ray source in the field and is probably confused. In addition, the re-observed spectrum of QSF3:22, previously classed as a possible QSO, clearly reveals this object to be a star. Thus, only four of the 16 UVX QSOs observable in practice with Rosat remain undetected at the X-ray fluxes reached here.

In contrast, none of the narrow-emission-line galaxies identified in the UVX survey were detected by Rosat.

To identify the optical counterparts of the 27 remaining X-ray sources, spectroscopic observations were made with the Autofib fibre optic system¹⁷ at the Anglo-Australian Telescope (AAT). The field of view of Autofib is $40'$ in diameter and is therefore well matched to the central PSPC field. We then used a transformation based on the 12 UVX QSOs detected by Rosat to cross-correlate the corrected X-ray positions with the COSMOS $1''$ positions for stars and galaxies on the IIIaJ plate. For 22 out of the 27 additional sources, an optical counterpart with $B < 22.3 \text{ mag}$ (the limit for a reliable detection on the UK Schmidt plate) was found within $15''$ of the transformed X-ray position. Of the 34 sources with optical counterparts, five had two and three had three counterparts within $15''$. In these cases, positioning restrictions with the Autofib system meant that it was possible to place a fibre only on the object nearest to the position of the X-ray source. It is unlikely, however, that this restriction has led to source misidentification in this survey, as in six cases the closest object has been identified as a QSO or a probable QSO. In a further case (XSF3:78) the alternative object was observed later. In the one remaining case (XSF3:67) the alternative object shows excess ultraviolet flux and is likely to be the X-ray source.

The AAT observations were made during 14–16 November 1990, when an 18,000-s Autofib exposure was obtained, and on 11 December 1990, when a 7,000-s exposure service observation was made by D. Hatzidimitriou. During both these runs, Autofib was operated in conjunction with the Royal Greenwich Observatory spectrograph and the IPCS detector, giving a spectral resolution of 12 Å over the wavelength range $3,400 < \lambda < 7,200 \text{ Å}$. All the new X-ray sources were observed, and nine out of the 12 UVX QSOs were re-observed. The spectroscopic identifications are summarized in Table 1. Thirteen new QSOs (emission line full width at half maximum $> 1,000 \text{ km s}^{-1}$) have been found (including three probable QSOs), making a total of 25 QSOs detected by Rosat in the $40'$ diameter area. Spectra for the newly confirmed Rosat QSOs are presented in Fig. 2; spectra for the previously known QSOs have been presented elsewhere¹⁴. The new QSOs include a remarkable pair (XSF3:65), separated by only $5''$ on the sky but with different redshifts; $z = 1.586$ and $z = 2.077$. Although detected as a single QSO in the Autofib

TABLE 1 Rosat identifications

IAU name	Survey name	$S(0.1-2.4 \text{ keV})$ ($\text{erg cm}^{-2} \text{ s}^{-1}$)	Detection significance	B (mag)	Spectroscopic identification	z	N_{obj}	Comments
RX J0342.2-4350	XSF3:17	0.31×10^{-13}	6.5	18.4	Star	—	1	Composite spectrum
RX J0341.5-4351	XSF3:19	0.53×10^{-13}	10.6	—	?	—	0	
RX J0342.6-4353	XSF3:20	0.47×10^{-13}	11.9	19.7	QSO	0.564	1	
RX J0341.8-4353	XSF3:21	0.60×10^{-12}	153.9	<11.0	Star	—	1	Spectral type F
RX J0342.3-4355	XSF3:23	0.36×10^{-13}	10.0	22.1	QSO	2.551	1	
RX J0343.3-4355*	XSF3:24	0.42×10^{-13}	9.4	20.3	?	—	1	Non-stellar image
RX J0343.0-4355	XSF3:25	0.12×10^{-12}	30.5	20.8	QSO	0.867	1	
RX J0341.0-4357	XSF3:28	0.20×10^{-13}	4.5	22.0	?	—	1	QSF3:01
RX J0341.2-4357	XSF3:29	0.41×10^{-13}	10.8	22.3	QSO?	0.66†	3	
RX J0340.9-4402	XSF3:32	0.49×10^{-13}	13.1	20.8	QSO	1.521	1	QSF3:29
RX J0341.2-4402	XSF3:33	0.18×10^{-13}	5.3	—	?	—	0	
RX J0340.6-4403	XSF3:35	0.20×10^{-13}	3.8‡	—	?	—	0	QSF3:13
RX J0342.0-4403	XSF3:36	0.29×10^{-12}	113.6	19.2	QSO	0.635	1	
RX J0343.8-4403	XSF3:37	0.91×10^{-14}	1.7‡	21.1	?	—	1	Star + galaxy
RX J0342.6-4404	XSF3:38	0.30×10^{-12}	118.6	20.4	QSO	0.377	1	
RX J0341.2-4405	XSF3:42	0.24×10^{-13}	7.6	21.9	QSO	2.277	3	QSF3:40
RX J0343.4-4406	XSF3:44	0.24×10^{-13}	6.8	20.4	QSO	1.950	1	
RX J0341.0-4407	XSF3:45	0.57×10^{-13}	16.7	21.1	QSO?	1.76†	1	QSF3:20
RX J0342.5-4407	XSF3:46	0.16×10^{-13}	7.7	—	?	—	0	
RX J0342.0-4409	XSF3:48	0.13×10^{-13}	6.3	21.3	QSO	1.828	1	QSF3:27
RX J0341.3-4410	XSF3:51	0.33×10^{-13}	10.6	21.6	Star?	—	1	
RX J0341.6-4411	XSF3:53	0.29×10^{-13}	10.5	20.4	QSO	1.897	3	UVX
RX J0341.0-4412	XSF3:54	0.67×10^{-13}	19.0	21.6	QSO	1.808	1	
RX J0343.9-4411	XSF3:56	0.13×10^{-13}	2.2‡	22.0	QSO?	2.21†	2	QSF3:32
RX J0342.3-4412	XSF3:57	0.42×10^{-13}	15.9	21.8	QSO	1.091	1	
RX J0341.9-4414	XSF3:58	0.40×10^{-13}	14.6	20.1	QSO	1.478	1	QSF3:32
RX J0341.5-4414	XSF3:59	0.14×10^{-13}	4.5	22.2	QSO	1.09†	1	
RX J0341.1-4415	XSF3:61	0.24×10^{-13}	6.5	21.9	QSO	1.73†	1	Elliptical galaxy
RX J0340.9-4415	XSF3:63	0.33×10^{-13}	7.6	19.4	Galaxy	0.180	1	
RX J0343.0-4416	XSF3:64	0.56×10^{-13}	16.8	<11.0	Star	—	1	Spectral type F
RX J0342.2-4416	XSF3:65	0.40×10^{-13}	13.6	19.2	QSO	2.077	2	
	XSF3:65a				QSO	1.586		Double QSO
RX J0341.9-4416	XSF3:66	0.33×10^{-13}	11.5	19.5	QSO	1.797	2	
RX J0341.1-4417	XSF3:67	0.38×10^{-13}	9.7	22.3	?	—	2	QSF3:31
RX J0342.4-4417	XSF3:68	0.25×10^{-13}	7.9	—	?	—	0	
RX J0343.2-4420	XSF3:70	0.45×10^{-13}	10.7	21.2	QSO	1.378	1	QSF3:45
RX J0341.9-4422	XSF3:71	0.70×10^{-13}	18.5	21.5	QSO	1.08†	1	
RX J0342.9-4422	XSF3:72	0.97×10^{-13}	23.3	19.8	QSO	0.897	1	QSF3:47
RX J0341.4-4425	XSF3:77	0.63×10^{-13}	11.5	19.5	QSO	0.794	1	
RX J0342.0-4425	XSF3:78	0.88×10^{-14}	1.9‡	20.6	?	—	?	Non-stellar image
	XSF3:78a			18.4	Star	—		

* Uncertain optical position.

† Uncertain redshift.

‡ $>4\sigma$ detection in hard band.

observations, subsequent conventional slit spectroscopy at the AAT enabled both QSOs to be identified. It is possible that the galaxy or cluster of galaxies in which the QSO with the lower redshift lies is lensing the one at higher redshift. Only two of the newly identified QSOs are brighter than the UVX survey limit. The double QSO, XSF3:65, was missed by the UVX survey because it was classified by COSMOS as a galaxy. XSF3:20 ($z=0.564$) lies in the redshift range where the $U-B$ colours of QSOs show significantly less ultraviolet excess, and failed by only $\Delta(U-B)=0.03$ mag to be included in the original UVX sample. The UVX technique thus misses $\sim 10\%$ of possible sources, confirming previous incompleteness estimates¹⁸.

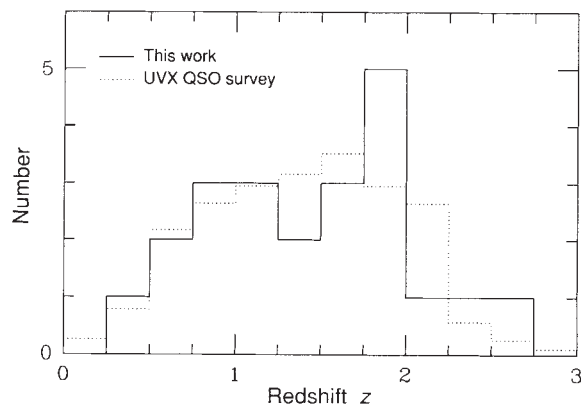
Apart from the QSOs, we have identified three definite stars. Two of the stars have $B < 11$ mag and are classed as F stars. XSF3:17 is classed as a composite spectrum object, comprising a hot O/B star spectrum at $\lambda < 4,000 \text{ \AA}$ and a cooler G star spectrum at $\lambda > 4,000 \text{ \AA}$. The spectrum for one of these stars, XSF3:21, is also shown in Fig. 2. A further two possible X-ray stars were detected, one (XSF3:78a) as an alternative optical identification for an X-ray source, and the other (XSF3:51) at low signal-to-noise ratio.

With the exception of the double QSO, optical counterparts of three X-ray sources were identified by COSMOS as non-stellar

images. XSF3:37 has also turned out to be a double image, this time a star-galaxy pair with $2.5''$ separation, from an AAT plate of the field taken in December 1990 by D. Malin. The fibre spectrum of the merged object is unusually blue but shows no emission lines. XSF3:63 has $z=0.180$ and is clearly identified as an elliptical galaxy on the AAT photograph; its spectrum is also shown in Fig. 2. Finally, XSF3:24 is a galaxy with a stellar nucleus. Unfortunately its optical position may be affected by the diffraction spike of a neighbouring star, which is close to it on the plate, and it remains unidentified. In all, eight sources were unidentified spectroscopically; these include the five X-ray sources with no optical counterpart at $B < 22.3$ mag. Apart from XSF3:24, the remaining two objects (XSF3:28, XSF3:67) have $B \geq 22$ mag and their spectra are poor.

Based on the 24 QSOs (counting the double QSO, XSF3:65, only once) identified in the central $40'$ field, the X-ray QSO surface density obtained in this survey is $69 \pm 14 \text{ deg}^{-2}$. At the completeness limit of $S(0.5-2.0 \text{ keV}) = 1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$, the QSO density is $63 \pm 13 \text{ deg}^{-2}$. This should be compared with the QSO surface density of $32 \pm 9 \text{ deg}^{-2}$ at $2.5 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ found in the Pavo field¹⁰. The X-ray QSO number-redshift relation is shown in Fig. 3. The median QSO redshift for this survey is $z=1.52$, higher than the values of

FIG. 3 The number-redshift relation for the 22 X-ray-selected, confirmed QSOs. The $n(z)$ relation extends to larger redshifts than that for brighter X-ray-selected samples, with the highest-redshift QSO having $z=2.5$. The Rosat QSO $n(z)$ is more similar to $n(z)$ for the optically selected UVX QSO¹⁴ shown here as the dashed line.



$z=0.9$ derived for the Pavo deep survey¹⁰ and $z=0.3$ for the EMSS¹⁶. But the redshift distribution is very similar to that for the UVX QSOs identified in faint optical surveys¹⁴ (see Fig. 3). This emphasizes the degree to which these two samples overlap at the flux levels reached here.

X-ray source counts

With the high number of discrete sources detected in our survey, we can construct a $\log N : \log S$ relationship between number N , and flux, S . To ease comparisons with previous results, and to minimize effects due to intrinsic absorption and soft excesses in the source spectra, we base the X-ray count on source fluxes measured in the PSPC 0.5–2 keV band.

In Fig. 4, we show the derived $\log N : \log S$ diagram, for the non-stellar subsample of our data for sources in the 0.5–2 keV band, and for comparison, the $\log N : \log S$ derived from the Einstein Observatory EMSS¹⁶ and Deep¹¹ Survey (after correcting these to our passband). We note that there is good agreement at the higher fluxes with the Einstein results. Fitting a power-law model to the extragalactic sources with fluxes above $1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ yields a best-fit slope of -1.6 ± 0.3 ; this estimate takes into account a small bias¹⁹ due to increasing flux errors near our completeness limit. Fixing the slope to the euclidean value of -1.5 gives a normalization of $3.0 \pm 0.5 \times 10^{-16} \text{ sources sr}^{-1}$, which is close to the Einstein Deep Survey result¹¹ of $3.4 \pm 0.8 \times 10^{-16} \text{ sources sr}^{-1}$.

Furthermore, extrapolation of the EMSS $\log N : \log S$ with its -1.5 slope to our $1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ completeness limit predicts an extragalactic source density of 61 deg^{-2} , in reasonable agreement without our observed extragalactic source density of $89 \pm 16 \text{ deg}^{-2}$. At fainter fluxes the $\log N : \log S$ curve shows a sharp break, but as this feature occurs so close to our

flux limit, incompleteness, flux errors and confusion noise may be a problem. Hasinger *et al.*²⁰ have reported a similar feature in a comparable Rosat exposure at the north ecliptic pole (NEP). Certainly, if the discrete source count is not to saturate the Rosat soft X-ray background, then the $\log N : \log S$ slope must flatten before $1 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$.

A further constraint on the surface density of sources can be obtained by analysing the fluctuations in the PSPC background. To do this we have divided the central region of the field into $1'$ bins, rejected those with significant sources and then taken alternate bins, thus creating two subsets of independent samples. Trial $\log N : \log S$ distributions consisting of single power laws were then convolved with the instrument response and a predicted $P(d)$ curve calculated using standard techniques^{21,22}. These are fitted to the observed distribution, and the $\pm 1\sigma$ limits on $\log N : \log S$ are also shown in Fig. 4. We have constrained the normalization at $10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ to be within the range allowed by the source counts; the model counts are also cut at the faint flux limit where they would saturate the X-ray background. The best-fit value for the slope is -1.2 ± 0.15 , significantly flatter than the slope of the source counts at higher fluxes. We have also investigated to what depth the $\log N : \log S$ can continue with a euclidean slope. An extension below $\sim 3 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$ is ruled out at the 95% level even if there are no other discrete sources. This is a much more stringent limit than that determined from similar analyses of Einstein data^{22–24}.

In Fig. 4 we also show the QSO $\log N : \log S$ predicted from the EMSS results for the evolution of luminosity functions of QSOs and active galactic nuclei (AGN)²⁵. The prediction shows similar behaviour to that found from the $P(d)$ analysis, with the number count showing a gradual flattening from the euclidean beyond our source detection limit. A turnover in the

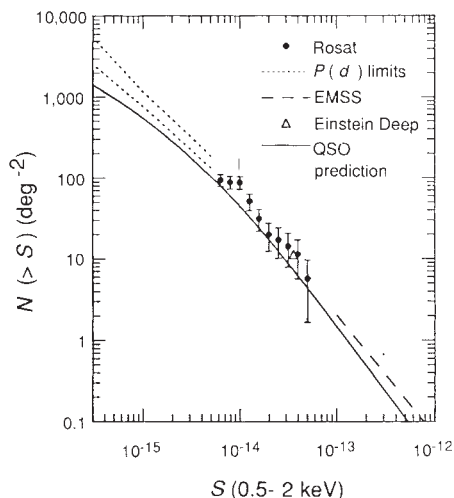


FIG. 4 The Rosat integral $\log N : \log S$ relation for all non-stellar sources in the 0.5–2 keV band, compared with the results from the Extended Medium Sensitivity Survey (dashed line) and the Einstein Deep Survey (triangle). The vertical line represents the completeness limit of the Rosat survey. The $\pm 1\sigma$ limits on the $\log N : \log S$ at fainter fluxes from $P(d)$ analysis of the background fluctuations are shown as dotted lines. The QSO and AGN number count predicted from the EMSS luminosity function evolution parameters is represented as the solid line.

$\log N : \log S$ relation is also expected from predictions derived from the QSO optical luminosity function²⁶. By a flux of $\sim 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$, however, the EMSS prediction falls below the limits of the $P(d)$ analysis. This could be evidence for a new population of sources. But the QSO count could be higher than predicted if a slightly higher rate of evolution were adopted or if the slope of the faint end of the QSO and AGN luminosity function is significantly steeper at $z > 0.5$. Alternatively, the lower bound from the $P(d)$ analysis may be uncertain if some of the observed variance is due to instrumental effects. Thus the fact that the prediction falls lower than the $P(d)$ limits cannot immediately be taken to indicate a new source population.

We conclude that the upper limit from the $P(d)$ analysis implies at least a gradual flattening in the observed $\log N : \log S$ for sources at faint X-ray fluxes, as would be predicted from the evolution of the QSO X-ray luminosity function.

Contribution to the X-ray background

Our data provide strong constraints on the contribution of point sources to the X-ray background. A simple, model-independent estimate of the lower limit to the contribution comes from taking the ratio of integrated source counts above 0.5 keV (to minimize the diffuse contribution from our Galaxy) to the total number of counts (including sources) detected in the inner 20' radius of the Rosat field. This gives a value of $\sim 35\%$. Excluding the stellar sources from the calculation gives a value of $30 \pm 5\%$, most of which ($\sim 65\%$) comes from the identified QSOs.

The flattening in the $\log N : \log S$ slope from $S^{-1.5}$ to $S^{-1.2}$ increases the source density required to produce 90% of the background from $2,000 \text{ deg}^{-2}$ integrating to $1 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$, to $4,000 \text{ deg}^{-2}$ integrating to $5 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$. The increase in source density required by the flatter slope is therefore only a factor of 2 and it is thus still possible that discrete sources account for almost all of the 0.5–2 keV X-ray background. The EMSS QSO luminosity function model predicts that 35–40% of the background is due to QSOs²⁵. But as noted above, uncertainties in the model could raise (or lower) this contribution. For example, the model currently underpredicts the total Rosat source density slightly at $1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ (see Fig. 4). If details of the evolution model were adjusted to improve this agreement, this could double the predicted QSO contribution to the background.

Consideration of the detected source (QSO) and background X-ray spectra may shed further light on the question of whether sources other than AGN are required to explain the Rosat X-ray background. First, we note that there is some evidence for an upturn in the X-ray background spectrum in the 0.5–2 keV band. If the extragalactic spectrum measured at higher energies²⁷ is extrapolated into the Rosat passband, then only $\sim 60\%$ of the Rosat background counts are contributed by the extragalactic component. A similar upturn has been claimed in previous X-ray

observations of the soft X-ray background²⁸. We emphasize that the sources represent 35% of the total Rosat 0.5–2 keV background, including the excess.

The resolution of the PSPC allows us to estimate the detected source contribution to the background as a function of energy. In Fig. 5 we show the summed spectra of the QSO sources and of the background. As the hardness ratios of the faint, unidentified sources are mostly consistent with their being QSOs, they have been included in the summed QSO spectrum. The source and background spectra from the NEP field of Hasinger *et al.*²⁰ are also plotted. With the exception of points at energies less than 0.3 keV (see below), these are in reasonable agreement with our data.

First, it is clear that the detected QSO contribution to the background reaches a peak at $\sim 1 \text{ keV}$, where it is $\sim 35\%$. In the range 1.0–2.0 keV, power-law fits to the source spectrum suggest an energy index, $a_x = 1.3 \pm 0.2$, whereas for the background we find $a_x = 0.7 \pm 0.2$. This result differs from that of Hasinger *et al.*²⁰, who found $a_x = 1.3$ for the sources but $a_x = 1.2$ for the background in the NEP field. Unfortunately, because of the limited energy range for fitting, the small differences in the 1–2 keV background spectra seen in Fig. 5 can lead to large variations in the spectral slopes obtained. Given the uncertainty, adopting a value of $a_x = 1.0 \pm 0.3$ for the average background slope seems reasonable. Despite the error, the background spectral slope still seems steeper than the slope of $a_x = 0.4$ found at harder energies, in agreement with the 40% excess in the 0.5–2 keV band noted above. Although our results also suggest that the source spectrum is steeper than the background, we cannot reject the possibility that the spectral indices are the same.

At lower energies (0.25 keV), the fraction of the total Rosat background explained by detected sources drops considerably, to less than 10%. This behaviour is expected because of an increasing galactic component at soft energies. We also note that the QSO source intensity at 0.25 keV is a factor of 4 lower in the NEP data than in ours, when normalized at energies in the range 1–2 keV. The depletion of the QSO spectrum in the NEP field is reasonably consistent with H I absorption: the H I column in the NEP field is $4 \times 10^{20} \text{ cm}^{-2}$, whereas in QSF3 it is $1.7 \times 10^{20} \text{ cm}^{-2}$, leading to a relative extinction at 0.25 keV of a factor of 5. Further, if the $\sim 50\%$ depletion of the total 0.25 keV background in the NEP field relative to QSF3 is simply due to this absorption of the extragalactic component, then this leads directly to an estimate of $\sim 50\%$ for the total extragalactic component of the QSF3 background at 0.25 keV. But this estimate relies on the still unverified assumption that the size of the galactic component in each field is the same.

Finally, we return to discuss the possible background components in the 1–2-keV range. As the intensity of the 1–2-keV background is very similar in our QSF3 field at $b^{\text{II}} = -50^\circ$ and the NEP field at $b^{\text{II}} = 30^\circ$, we shall assume that this indicates

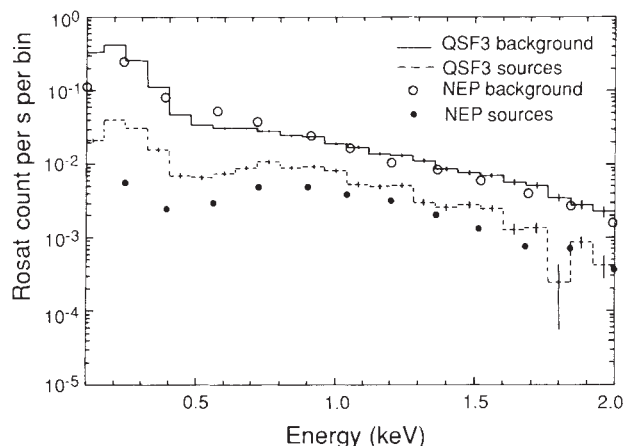


FIG. 5 The total non-stellar source spectrum compared with the background spectrum, including sources, in the 0.1–2.4-keV band for the QSF3 field. Also shown for comparison are the same results for the north ecliptic pole field of Hasinger *et al.*²⁰. There are only ~ 6 independent resolution elements across these spectra.

that the extragalactic component dominates any galactic component. A very local, isotropic galactic component, such as is believed to contribute at 0.25 keV, however, may still not be completely ruled out.

The observed steepening of the background spectral slope relative to energies above 3 keV could suggest that the background in the Rosat passband is dominated by QSOs with a spectral index $\alpha_x \approx 1$. Also, our fluctuation analysis results allow >90% of the background to be produced by sources with a sky density of 4,000 deg⁻². Uncertainties in QSO X-ray evolution and spectral parameters may allow so high a density of QSOs and AGN; thus it seems that our number counts and X-ray spectra still permit almost all of the X-ray background in the 1–2-keV range to be produced by QSOs. At harder energies, the canonical, steep ($\alpha_x = 0.7$ –1) values for QSO as Seyfert spectral indices makes it less easy for currently known AGN populations to contribute significantly to the hard background. Ginga spectra in the range 2–20 keV for Seyfert galaxies are complicated and show evidence of flat components²⁹. Information on the 2–20-keV spectra of high-luminosity QSOs is more limited, but there is little evidence for additional components or spectral curvature³⁰ and the mean QSO spectral index is ~ 0.8 . It is therefore unlikely that QSOs dominate the hard X-ray background. This possibility is still not excluded, however, at softer energies.

If a second component is required to explain the hard X-ray background spectrum then it may also make a significant contribution to the 1-keV background. Models in which QSOs with $\alpha_x \approx 1$ account for roughly half the background at 1 keV, and another component with $\alpha_x \approx 0.3$ in the range 0.5–10 keV accounts for the other half, also give reasonable fits to the overall background spectrum.

This second component (of which we have no detections) could be a truly different source population such as starburst galaxies⁹; alternatively, it could be due to a population of AGN with different spectral properties, or spectral properties that depend on energy³¹ or evolutionary state³². Possible sources suggested on the basis of the Ginga spectra at 2–20 keV include highly absorbed low-luminosity AGN³³ (Seyfert II and narrow-emission-line galaxies) and AGN dominated by Compton reflection³⁴. We note that our spectral results suggest that at redshifts of ~ 1 –2 the spectral index of high-luminosity QSOs is not substantially different from similar nearby objects. The high

Rosat detection rate of the UVX QSOs also tends to exclude the possibility that objects of high optical luminosity form the alternative population.

There is therefore no doubt that QSOs account for at least one-third of the X-ray background at 1 keV. But the question of whether they alone are the dominant contributor or whether a new, faint, source component is required remains finely balanced. A resolution of this question will require new data. For example, further optical spectroscopic follow-up of our unidentified X-ray sources is required to check for any significant new population of faint sources. Searches for spatial correlations between X-ray background fluctuations and optical or radio sources in deep images of the QSF3 field may also produce constraints on new source populations at limits inaccessible to optical spectroscopy. Information is also required on the 0.1–10-keV spectra of sources with X-ray fluxes of 10^{-15} erg cm⁻² s⁻¹ and this will be obtainable in the near future with instruments such as JET-X.

Conclusions

In a very deep exposure with the Rosat PSPC detector, we have detected 39 X-ray sources in an area of 0.35 deg². Follow-up spectroscopy at the Anglo-Australian Telescope has identified 24 QSOs amongst these sources for a QSO surface density of 69 ± 15 deg⁻². We find that 75% of QSOs with $B < 21.2$ mag are X-ray emitters and have established that faint, X-ray surveys are dominated by QSOs of $z \approx 1.5$. The Rosat background intensity at 1 keV shows an excess of $\sim 40\%$ over the extrapolated 3–10-keV spectrum, indicating an upturn of the background spectrum in the Rosat band. The PSPC 0.5–2-keV background fluctuations show a flattening of the soft X-ray log N :log S curve beyond $S(0.5$ –2 keV) $\approx 1 \times 10^{-14}$ erg cm⁻² s⁻¹, but discrete sources are still capable of supplying the total Rosat background at 1 keV. Direct comparison of the total QSO spectrum with the background spectrum gives a lower limit of 30% for the QSO contribution to the cosmic X-ray background at 1 keV. Another extragalactic component with a flatter spectrum may also be required to explain a significant proportion of the background at 1 keV and a larger fraction at harder energies. But on the basis of the X-ray spectra and number counts, there is still the possibility that QSOs account for almost all of the X-ray background at 1 keV. $\square$

Received 3 June; accepted 19 August 1991.

- Setti, G. & Woltjer, L. *Astr. Astrophys.* **76**, L1–L2 (1979).
- Tananbaum, H. et al. *Astrophys. J.* **234**, L9–L13, (1979).
- Schmidt, M. & Green, R. F. *Astrophys. J.* **305**, 68–82 (1986).
- Barcons, X. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **230**, 189–206 (1988).
- Marshall, H. L. et al. *Astrophys. J.* **283**, 50–58 (1984).
- Georgantopoulos, I. & Shanks, T. in *X-Ray Astronomy: Proc. 23rd ESLAB Symp.* (eds Hunt J. & Batrick, B.) 939–943. (ESA Publications Division, 1989).
- Guilbert & Fabian, A. C. *Mon. Not. R. astr. Soc.* **220**, 439–451 (1986).
- Schaeffer, R. P. & Silk, J. *Astrophys. J.* **333**, 509–521 (1988).
- Griffiths, R. E. & Padovani, P. *Astrophys. J.* **360**, 483–489 (1990).
- Griffiths, R. E. et al. in *Proc. Workshop on Optical Surveys for Quasars* (eds Osmer, P. S. et al.) 351–360 (Astronomical Society of the Pacific, San Francisco: 1988).
- Primini, F. A. et al. *Astrophys. J.* (1991). (in the press).
- Briel, U. G. et al. *Proc. SPIE* (in the press).
- Pfeffermann, E. et al. *Proc. SPIE* **733**, 519–525 (1986).
- Boyle, B. J., Fong, R., Shanks, T. & Peterson, B. A. *Mon. Not. R. astr. Soc.* **243**, 1–56 (1990).
- Boyle, B. J., Jones, L. R. & Shanks, T. *Mon. Not. R. astr. Soc.* **251**, 482–507 (1991).
- Gioia, I. M. et al. *Astrophys. J. suppl. Ser.* **72**, 567–619 (1990).
- Parry, I. R. & Sharples, R. M. in *Fibre Optics in Astronomy* (ed. Barden, S. C.) 93–98 (Astronomical Society of the Pacific, San Francisco: 1988).
- Boyle, B. J. thesis, Univ. of Durham (1986).

- Murdoch, H. S., Crawford, D. F. & Jauncey, D. L. *Astrophys. J.* **183**, 1–13, (1973).
- Hasinger, G., Schmidt, M. & Trumper, J. *Astr. Astrophys.* **246**, L2–L5 (1991).
- Condon, J. J. *Astrophys. J.* **188**, 279–286 (1974).
- Barcons, X. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **243**, 366–371 (1990).
- Hamilton, T. T. & Helfand, D. J. *Astrophys. J.* **318**, 93–102 (1987).
- Soltan, A. M. *Mon. Not. R. astr. Soc.* **250**, 241–252 (1991).
- Maccacaro, T. et al. *Astrophys. J.* **374**, 117–133 (1991).
- Georgantopoulos, I. thesis, Univ. of Durham (1991).
- Marshall, F. E. et al. *Astrophys. J.* **235**, 4–10 (1980).
- McCammon, D. & Sanders, W. T. *Ann. Rev. Astr. Astrophys.* **28**, 657–688 (1990).
- Pounds, K. A., Nandra, K., Stewart, G. C., George, I. M. & Fabian, A. C. *Nature* **344**, 132–133 (1990).
- Williams, O. R. et al. *Astrophys. J.* (in the press).
- Schwartz, D. A. & Tucker, W. H. *Astrophys. J.* **332**, 157–162 (1988).
- Boldt, E. & Leiter, D. *Astrophys. J.* **276**, 427–433 (1982).
- Warwick, R. S. & Stewart, G. C. in *X-Ray Astronomy: Proc. 23rd ESLAB Symp.* (eds Hunt J. & Batrick, B.) 727–731 (ESA Publications Division, 1989).
- Fabian, A. C., George, I. M., Miyoshi, S. & Rees, M. J. *Mon. Not. R. astr. Soc.* **242**, 14P–16P (1990).

ACKNOWLEDGEMENTS. We thank the Director and staff of the Anglo-Australian Telescope for assistance with the Autofib observations and the Rosat staff for support with the X-ray observations and analysis. T.S. thanks the Royal Society and B.J.B. thanks the SERC and Institute of Astronomy, Cambridge for support. G.C.S. and R.E.G. acknowledge the support of the SERC and NASA, respectively.

The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation

D. R. Madden^{*†}, J. C. Gorga^{†‡}, J. L. Strominger[†] & D. C. Wiley^{†§}

^{*} Committee on Higher Degrees in Biophysics, [†] Department of Biochemistry and Molecular Biology, [§] Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

X-ray crystallography reveals electron density in the antigen-binding site of HLA-B27 that is an interpretable image of nonameric peptides in a largely extended conformation. Clear density exists for the main chain and several side chains and is consistent with the sequence of 11 nonameric self-peptides eluted from HLA-B27 (see accompanying article¹). Pockets in the antigen-binding cleft bind four side chains and the amino and carboxyl termini of the peptide.

CYTOTOXIC events in the cellular immune response begin with specific cell-to-cell recognition. The hypervariable immunoglobulin-like T-cell receptor on the surface of cytotoxic T cells (CTL) binds to the complex between a short peptide processed from a protein antigen and a target cell's class I major histocompatibility glycoprotein (reviewed in ref. 2). The range of peptides that can be presented for immune surveillance is enhanced by the coexpression of several different class I histocompatibility antigens on the cell, and, more importantly, by the ability of each histocompatibility molecule to bind many peptides of diverse sequences. In the population the major histocompatibility complex (MHC) antigens are highly polymorphic, having been selected to increase the diversity of peptides that can be presented to CTL (refs 3, 4).

The three-dimensional structures of the human class I histocompatibility antigens HLA-A2 and HLA-Aw68 revealed a peptide-binding groove lined with polymorphic residues^{3,5-7}. Electron density bound in the groove, although not definitively interpretable, seemed to represent short peptides bound to the crystalline HLA (refs 3, 5, 6), and a polyalanine 9-mer was modelled into the density associated with HLA-A2 (ref. 7; M. A. Saper, unpublished data). The inability to dissociate the endogenous mixture of peptides and readily prepare molecules with one peptide bound⁸ has made it difficult to determine either how one peptide binds to a given site or how a given site can accommodate any one of a large number of peptides with different sequences. Indeed, peptide binding seems to be associated with the folding and assembly of class I MHC molecules⁹⁻¹³ and only strategies involving reconstitution or *de novo* biosynthesis of class I molecules in the presence of single peptides will yield stoichiometric complexes^{12,15,38-40}.

Here we examine the unusually clear electron density in the binding site of HLA-B27. It indicates that the amino-acid sequence of bound nonamer peptides would be restricted at four positions (P2, P3, P7, P9; where P_n is the *n*th position in the peptide) to a B27-specific sequence motif, because these peptide side chains are bound in 'pockets' in the bottom and

sides of the binding groove. To a varying degree, the pockets can accommodate several different peptide side chains, consistent with evidence from peptide sequences¹. A T-cell receptor could interact directly with the four or five peptide side chains (P1, 4, 5, 6, 8) that appear accessible and could perhaps indirectly sense amino-acid substitutions at the more buried positions through shifts in the peptide main chain. The main chain of the first two and last two peptide amino acids and the peptide's charged termini seem to bind to conserved polar atoms at both ends of the binding groove.

Structural determination

Because HLA-B*2705 crystallized with two molecules in the crystallographic asymmetric unit¹⁴, an atomic model could be determined on the basis of the coordinates of HLA-A2 but free of bias from that structure. Iterative real-space phase averaging yielded high-quality electron density maps from starting models with one domain omitted at a time, and the model for that domain was built only into clear density (see Table 1 caption). The refined model was used to calculate new maps, which were again improved by phase averaging. Until the final stage, the model was extended only into clear electron density. Throughout refinement, both averaged and unaveraged maps were examined, and non-crystallographic symmetry constraints or restraints were applied. The final structure has a crystallographic *R* factor of 18.3% from 6.0–3.0 Å, and has reasonable geometric parameters (see Table 1). The structure of HLA-B27, the histocompatibility antigen associated with the disease ankylosing spondylitis (reviewed in ref. 17), is very similar to that of HLA-A2 and HLA-Aw68 (refs 5–7), as is to be expected from the 90% sequence identity. Details of the structure will be described after high resolution refinement (D.R.M. *et al.*, in preparation).

Electron density suggestive of a short peptide was observed in the peptide-binding groove in the earliest maps, as well as in the final refined maps (Fig. 1*a, b*). Presumably, it is the averaged image of a mixture of different bound peptides¹. To analyse the protein-peptide interaction, we built several peptide models and subjected them to refinement. But to avoid introducing bias into the electron density maps the peptide models were not included in the HLA-B27 refinement or in any of the electron density maps discussed here. The geometry of the current peptide model is standard: a data base search using the diagonal facility of FRODO (refs 18, 19) reveals a number of overlapping tetra- and pentapeptides with similar geometries.

Peptide conformation

The current model for a peptide in the cleft is a nonamer in a largely extended conformation, with side chains protruding in a roughly alternating pattern, as is found in extended β strands. There is a kink at amino acids three and four, such that the side chain of position 3 (P3) points forward along the direction of the peptide, fitting under the main chain of P5. The side chain of P4 points up (Fig. 1*c*). P1 and P2 are close to the floor of the cleft, with the P2 side chain in the '45 pocket' previously described^{6,7} (Fig. 2). There is no side chain density for P1, but a long side chain could reach solvent. The segment from P5 to

[‡] Present address: Department of Pediatrics, Children's Hospital of Pittsburgh, 6130 Rangos Research Building, 3705 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA.

FIG. 1 The peptide model and extra electron density in the peptide binding cleft of HLA-B27. *a*, Top view of the cleft, with the $\alpha 1$ α -helix at the top of the figure. Red, $C\alpha$ trace of the HLA-B27 binding groove helices, with Trp 147 side chain in yellow; yellow, peptide model; blue, $2F_o - F_c$ real-space phase-averaged electron density within an envelope describing the cleft, contour equals 1σ . *b*, Side view of the cleft, looking from the $\alpha 2$ α -helix (absent) towards the $\alpha 1$ α helix. *c*, Peptide model, viewed as in *b*.

METHODS. As an initial model of the bound peptide, a polyalanine octapeptide was built into the observed extra density with its N terminus at the left-hand (Y171) end of the cleft, and regularized using FRODO¹⁸. Hypothetical side chains were added at several positions, as suggested by the electron density and the binding environment. Another model with the reverse sequence was built into the density with its N terminus at the right-hand (Y84) end of the cleft. The two models were subjected to a simulated annealing slow-cooling protocol from 4,000 K to 300 K³⁷, and only the one with its N terminus at the left-hand end (Tyr 171) of the cleft achieved an acceptable peptide conformation. This peptide also fits the density somewhat better, with its penultimate carbonyl oxygen in a prominent electron density bulge and in a position to hydrogen-bond to the pyrrole nitrogen of the conserved Trp 147. An additional residue was added at the N-terminus, yielding a nonamer. This basic structure was modified to accommodate biochemical sequence information as it became available. The model peptide (sequence: ARYAASTEEL) is based on the histone peptide sequence eluted from HLA-B27 (ref. 1) with alanine at positions 1, 4, 5.

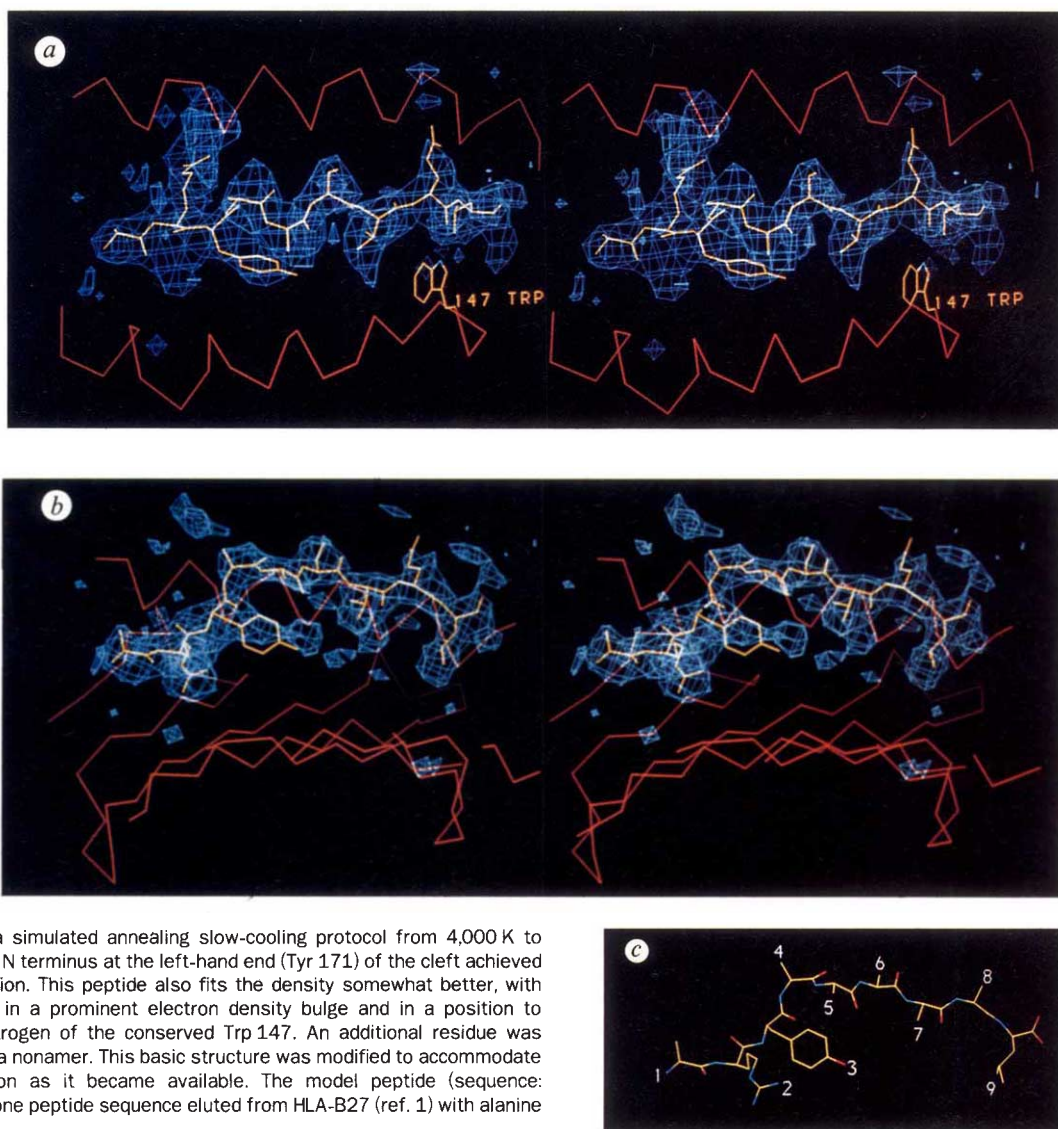


FIG. 2 Side-chain specificity. The charged and polar side chains of HLA-B27 in the '45 pocket', together with model of peptide side chain P2 (arginine, green) and electron density (blue, contoured as in Fig. 1).



TABLE 1 Data collection and refinement statistics

Resolution (Å)	Fraction complete	Number of independent reflections	Fraction $I > 3\sigma$	R_{merge}	Number of multiply-measured reflections	R_{cryst}
10.0–6.10	0.97	1,852	0.95	0.085	1,710	0.393
6.10–4.78	0.97	2,594	0.95	0.093	2,325	0.204
4.78–4.06	0.95	3,084	0.94	0.103	2,372	0.160
4.06–3.59	0.88	3,262	0.82	0.119	1,727	0.175
3.59–3.26	0.53	2,204	0.65	0.078	13	0.210
3.26–3.00	0.49	2,215	0.50	0.032	7	0.238
10–3 Å	0.75	15,211	0.81	0.098	8,154	0.218
6–3 Å	0.73	13,243	0.79	0.102	6,339	0.183

R.m.s. deviations in geometry: bond length, 0.017 Å; bond angle, 3.7°; dihedral angle, 25.9°. Data were collected from 5 crystals¹⁶ on a Xentronics area detector³¹, processed using BUDDHA programs³² and scaled with the CCP4 package (P. Evans, personal communication). Initial phases were determined by molecular replacement (refs 33–35; and P. Fitzgerald, personal communication). The molecules are related by an approximately twofold rotational symmetry axis (179.6° rotation, 3 Å translation), within 1° of the crystallographic a axis, as expected from pseudo-mm precession photographs and the existence of a nearly isomorphous P2 crystal¹⁶. To derive a model independent of the HLA-A2 structure, the four (~90-amino-acid) domains α_1 , α_2 , α_3 and β_2 -microglobulin were individually omitted from the model used to calculate phases for the electron density maps, which were each iteratively real-space phase-averaged to convergence³⁶. Each map yielded clear electron density for the omitted domain, and an unbiased partial model (69% complete) of the whole structure was constructed one domain at a time. Model was extended only into clear density, and its geometry regularized by FRODO¹⁸ and X-PLOR³⁷. The process was repeated until it yielded no new density. Subsequent refinement included least-squares positional and temperature factor refinement, as well as positional refinement by simulated annealing slow-cooling³⁷. New molecular envelopes and non-crystallographic symmetry relationships were determined. In the final stage, the resolution range for refinement was 6–3 Å, and strict non-crystallographic symmetry constraints were relaxed. Several further rounds of map averaging, building and refinement yielded the final structure. Eleven charged or polar surface sidechains were without significant electron density. No water molecules have been included in the model. $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$; $R_{\text{cryst}} = \sum |F_o - F_c| / \sum F_o$; I , measured diffraction intensity; σ , standard deviation on I ; F_o and F_c , observed and calculated structure factor amplitudes.

P9 is twisted so that P5 and P6 point roughly back and forth across the width of the binding site, with P5 towards the α_2 α -helix and P6 towards the α_1 α -helix, whereas side chains P7, P8, and P9 alternatively point approximately into, out of, and into the binding site (Fig. 1a, b).

Overall, P2, P3, P7 and P9 face inward to contact the HLA molecule, and the terminal carboxyl and amino groups fit into the ends of the groove (Fig. 1b, c); P4 and P8 point up towards the solvent; P5 and P6 could be simultaneously in contact with the HLA molecule and accessible to direct contact by a T-cell receptor. The side chain density is relatively short for both P5 and P6. In terms of the previously described pockets^{6,7} containing polymorphic residues in the antigen-binding site, peptide side chain P2 fits into pocket B (near 45), P3 into pocket D (near 156), P7 into pocket E (near 152), and P9 into pocket C (near 74 and 116). The N terminus and P1 are in pocket A, and the C terminus and part of P9 in pocket F. Although there is no electron density to suggest that longer peptides may be bound, model building demonstrates that it is possible to extend the peptide at either end. Longer peptides would presumably result in the loss of some hydrogen bonds to the charged terminal groups, as well as a salt bridge at the C terminus (see below).

Peptide specificity

The size and amino-acid composition of the sites binding P2, P3, P9, and to some extent P7, suggest restrictions on the side chains that could be accommodated at these positions in a peptide's sequence. The most obvious of these restrictions is at position P2, where clear electron density reaches deep into the largely hydrophobic pocket that ends near a negatively charged Glu 45, and Cys 67 (Fig. 2). This site would be complementary to a long positively charged peptide side chain. In fact, Jardetzky *et al.*¹ find arginine at P2 in all eleven self-peptides eluted from HLA-B27, a result consistent with the sequences of three viral peptides known to bind to B27 (refs 20, 21; K. C. Parker *et al.*, manuscript submitted).

Peptide side chain P3 binds in a predominantly non-polar site near Leu 156 which contains two aromatic side chains (Tyr 99 and Tyr 159), so that a hydrophobic or aromatic side chain would be expected on the peptide. A polar atom (from His 114) is accessible at the side of the pocket so that a tyrosine could also be accommodated. Most of the HLA-B27 specific

peptide sequences examined have hydrophobic or aromatic side chains at P3. Residues as large as Trp and as small as Ser are found¹. The current model accommodates even these extremes, although it is possible that the binding of water molecules and small adjustments of the peptide main chain will occur in some cases.

P7 and P9 fit into two pockets, which in HLA-B27 form a large cavity between Asp 74 and Trp 147 which has a charged and polar surface. The electron density for the side chain at P7 is extensive, but bifurcated, pointing towards His 114 and Asn 97, and therefore difficult to interpret. This suggests that different side chains at P7 may be accommodated by contacting different residues in HLA-B27. Seven different side chains are found at P7 in the 14 HLA-B27 restricted peptides at P7 (ref. 1). Determining exactly how the binding site can accommodate, for example, either an arginine or a glutamic acid at P7 or whether for some peptide sequences P7 might rotate to point out of the site, will require higher resolution and possibly structures with single peptides bound. At position P9, the side chain could reach deep into the site to contact polar and negatively charged residues Asp 116 or Asp 77, or if shorter, to contact non-polar atoms of nearby Thr 143, Trp 147, Leu 81 or Tyr 123. The electron density at P9 does not fill the pocket, so shorter polar or even hydrophobic side chains may also be present. This is consistent with the six residues found at P9 so far (K, R, L, Y, N, A) (ref. 1).

Side chain P1 seems likely to influence the position of the peptide main chain atoms. A small side chain at this position permits considerable rotation around the residue 1 psi angle. A large side chain at position P1 would probably extend towards the solvent, restricting the main chain conformation and resulting in the loss of some hydrogen bonds (see later). Both small and large side chains are found at P1 in the available B27-restricted peptide sequences¹.

The side chains at positions 4, 5, 6, and 8 are all oriented towards the solvent or else horizontally within the cleft, such that the side chains are potentially exposed to solvent. The sequences at P4, 5 and 8 show the expected wide range of side chains. P6, although exposed to solvent, shows only short electron density pointing between alanine 69 and the aliphatic chain of Lys 70, both on the α_1 α -helix, which may restrict the permissible sequences. Consistent with this, P6 is restricted to small

polar and non-polar residues in the known peptide sequences¹. Further sequences may, however, reveal more diversity if the side chain can bend to point out of the site.

From model building experiments in which different side chains are placed at the protein-binding positions of the peptide, it is possible to accommodate the observed diversity of sequence with at most minor modifications of the overall main chain conformation.

Peptide termini

Two clusters of amino acids conserved in essentially all class I histocompatibility antigen sequences, one at each end of the peptide-binding groove^{3,5,7}, are positioned to contact polar atoms of the main chain of the first two and last two peptide amino acids and the charged termini, features common to all peptides. The importance of conserved residues for recognizing common aspects of peptides, such as their main chain and termini, was anticipated^{3,7} and a role for tyrosines 7 and 171 and threonine 143 was demonstrated by site-directed mutagenesis (F. Latron *et al.*, manuscript in preparation).

At the carboxyl terminus of the peptide, the positively charged ϵ -amino group of lysine 146 and the hydroxyl groups of tyrosine 84 and threonine 143 are in positions to donate hydrogen bonds to the peptide carboxylate (Fig. 3b). Residues 146 and 84 are conserved, and residue 143 is threonine or serine so that the polar hydroxyl group is conserved. The carbonyl oxygen of the penultimate peptide amino acid P8 appears positioned to hydrogen bond to the pyrrole nitrogen of the conserved tryptophan 147 (Figs 1a and 3b).

As already discussed, the interactions with the amino terminus are likely to depend on the P1 side chain. Hydrogen bonds to conserved tyrosines 159 (from the P1 carbonyl oxygen) and 171 (from the amino-terminal nitrogen) are possibly independent of the nature of the P1 side chain (Fig. 3a). Given a small side chain, additional hydrogen bonds may be made to conserved

tyrosines 59 (from the peptide ammonium group) and 7 (from the P1 carbonyl oxygen). In addition, non-conserved arginine 62 (G, Q, L, E, in other known class I sequences) appears positioned to hydrogen-bond to the P2 carbonyl oxygen. Other polar non-conserved residues in the cleft include glutamate 63 (Q, N, I-once) and aspartate 77 (N, S), which are within 4.0 Å of the P2 and P9 amide nitrogens, respectively.

Preliminary analysis indicates that ~ 800 Å² of protein and 1,200 Å² of peptide solvent-accessible surface are sequestered by the interaction observed, as calculated by ACCESS (M. Handschumacher and F. M. Richards, personal communication). This is consistent with the formation of a tight complex.

Discussion

From X-ray crystallographic evidence, the peptides bound to HLA-B27 seem to be nonamers whose conformation is that of extended β strand, with a kink at positions 3 and 4. The quality of the electron density from a probably heterogeneous mixture of peptides suggests that peptides of diverse sequence are accommodated within a relatively narrow range of backbone conformations. All the geometric features of this conformation are standard, and have been observed within a database of protein structures determined at high resolution by X-ray crystallography¹⁹. The N-to-C orientation is consistent with that suggested for HLA-A2 (F. Latron *et al.*, manuscript submitted) and with that suggested for a peptide bound to the class II histocompatibility antigen HLA-DR7 (ref. 22) on the basis of experiments in which compensating changes in the histocompatibility antigen and peptide sequences were discovered. Models largely composed of extended structure have been proposed for peptides bound to H2-K^d (ref. 23) and to class II (refs 22, 24–25) based on binding experiments and model building.

The peptide model is consistent with the pattern of sequence restriction observed in naturally processed self-peptides eluted from HLA-B27 (ref. 1). The side chains at positions 2, 3, 7 and

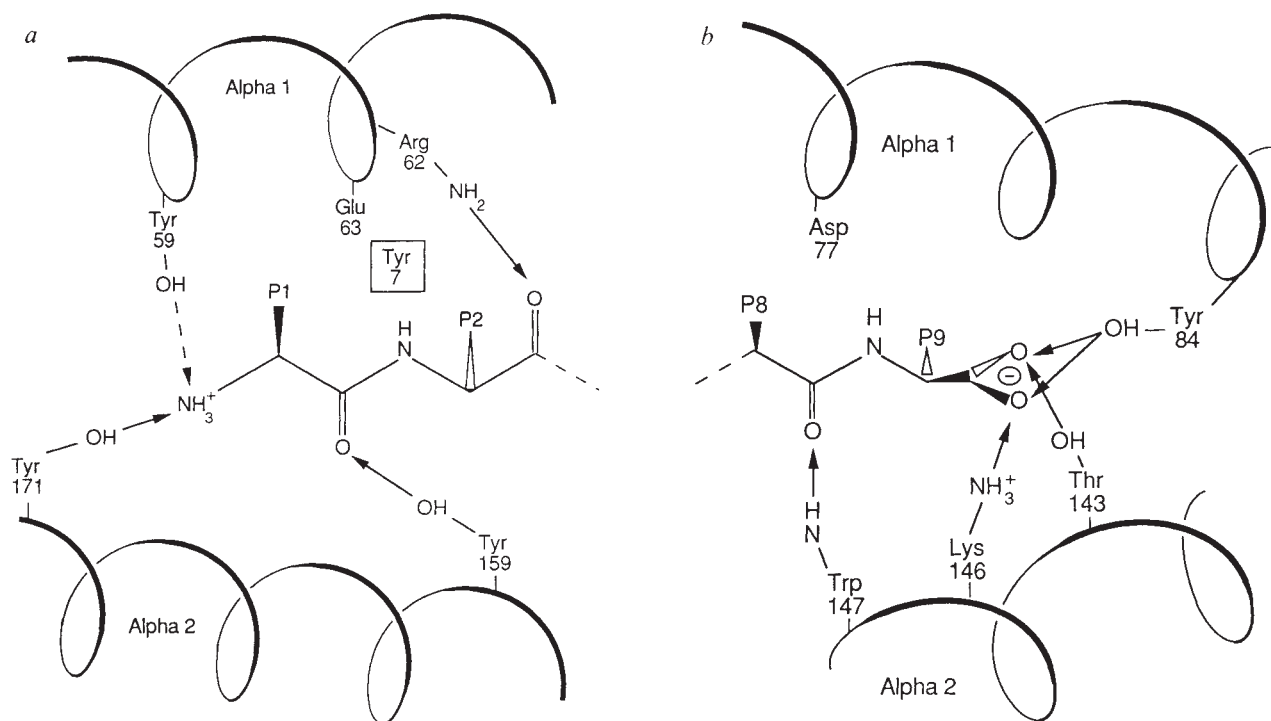


FIG. 3 Binding environment of the peptide termini. a, N-terminal and b, C-terminal two residues of the peptide model (stick figure). Residues of HLA-B27 currently within hydrogen-bonding distance of polar atoms in the model (solid forward arrow) and those potentially interacting with peptide

through water molecules or conformational shifts not ruled out by the density are labelled. Boxed, residue on the floor of the cleft; dashed arrow, hydrogen bond possible if P1 side chain is small.

9 are in close contact with the protein, in binding pockets whose geometry and chemistry determine an HLA-B27-specific sequence motif: arginine at 2, bulky non-polar or tyrosine and some small residues at 3, a mixture at 7, and at 9 Arg or Lys preferred but non-polar accepted¹. Sequence motifs have been suggested from binding data for class II molecules⁴¹ and from sequencing of the pool of bound peptide eluted from several class I alleles (ref. 26; G. M. Van Bleek and S. G. Nathenson, manuscript submitted). The particular stringency of binding at position P2 (14/14 Arg) coincides with the depth and narrowness of the '45 pocket'. The importance of a single amino acid is reminiscent of peptide binding to HLA-DR1, where a poly-alanine peptide containing one tyrosine binds as well as the viral peptide from which it was derived²⁷. Other pockets are more accommodating, binding any one of a small set of peptide side chains. P3, P7 and P9 bind in comparatively broad pockets, consistent with the large variation in residues found at these positions. P7 and P9 bind in a particularly wide cavity lined with polar and charged residues, and the density at P7 suggests that side chains may fit into more than one location within this pocket. In general, accommodation may also require the involvement of water molecules and in some cases movement of the peptide main chain which may be visible to the T-cell receptor. At P6 the electron density is too short to definitively explain the apparent sequence restriction at that position (see above). About one half of the bound peptide's side chains (P4, 5, 6, 8 and possibly 1) appear accessible enough to be contacted directly by a T-cell receptor and it is likely that the others, especially P3 and P7, could be recognized indirectly if they perturb the peptide backbone's location.

The polar atoms from conserved amino acids at the two ends of the binding site^{3,7} are here observed to contact the main chain polar atoms of the first two and the last two amino acids and the charged amino and carboxyl termini. The pockets at the ends of the class I MHC antigen-binding site therefore seem to be designed to recognize the ends of short peptides. This is consistent with the findings that peptides eluted from class I molecules are octamers and nonamers^{1,26,28-30}, and that the

natural short peptide eluted from H2-D^b forms a much more stable complex than peptides even one amino acid longer⁴⁰.

The mode of peptide binding seen here may have features general to other class I molecules, as the electron density in the binding sites of both HLA-A2 (refs 3, 5, 7) and HLA-Aw68 (ref. 6) is consistent with that described here (see Fig. 14 of ref. 7). Both showed peptide density deep in both ends of the binding groove and much more exposed at a bend or kink near the middle of the site. Anchor residues, positions of conserved sequence in peptides that bind a given MHC molecule, have been deduced from the sequencing of pools of peptides eluted from HLA-A2, H2-K^d, -K^b, and -D^b, and are found at positions P2 and P3 and at the C-terminal position, like the structure presented here. An anchor at P5 is observed in peptides eluted from H2-D^b and H2-K^b (ref. 26; G. M. Van Bleek and S. G. Nathenson, in preparation). In HLA-B27, P5 is relatively unconstrained by its contact with the HLA molecule. In other allelic variants, however, amino-acid differences in that region or a slightly altered conformation of the bound peptide could easily constrain the sequence at P5.

Our observations offer a general answer to the question of how one HLA molecule can form tight complexes with any one of a large number of peptides with diverse sequences. First, the MHC molecule determines the conformation of the bound peptide, and accommodates its side chains in binding-site pockets. Second, the peptide is only specifically recognized at four of its nine side chains, with further interactions involving main chain atoms at the two ends of the peptide and the charged termini. Third, some of the HLA pockets that recognize side chains can accommodate some diversity in peptide sequence. These strategies presumably evolved to increase the probability that at least one peptide expressed by an invading genome will bind a class I histocompatibility antigen and be presented to the immune system. For class II histocompatibility antigens, where the stimulus may be only a single protein antigen, peptide binding may be expected to involve more conserved contacts, fewer side-chain specificity pockets²⁷, or pockets with an even greater capacity for accommodating diversity. □

Received 22 July; accepted 14 August 1991.

- Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. *Nature* **353**, 326-329 (1991).
- Townsend, A. & Bodmer, H. A. *Rev. Immun.* **7**, 601-624 (1989).
- Bjorkman, P. J. et al. *Nature* **329**, 512-518 (1987).
- Parham, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4005 (1988).
- Bjorkman, P. J. et al. *Nature* **329**, 506-512 (1987).
- Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692-696 (1989).
- Saper, M. A., Bjorkman, P. J. & Wiley, D. C. *J. molec. Biol.* **219**, 277-319 (1991).
- Chen, B. P. & Parham, P. *Nature* **337**, 743-745 (1989).
- Townsend, A. R. M. et al. *Nature* **340**, 443-448 (1989).
- Townsend, A. R. M. et al. *Cell* **62**, 285-295 (1990).
- Schumacher, T. N. M. et al. *Cell* **62**, 563-567 (1990).
- Elliott, T. J. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5213-5217 (1990).
- Silver, M. L., Parker, K. C. & Wiley, D. C. *Nature* **350**, 619-622 (1991).
- Kvist, S. & Hamann, U. *Nature* **348**, 446-448 (1990).
- Benjamin, R. J., Madrigal, J. A. & Parham, P. *Nature* **351**, 74-77 (1991).
- Gorga, J. C., Madden, D. R., Prendergast, J. K., Wiley, D. C. & Strominger, J. L. *Proteins* (in the press).
- Khan, M. A. in *Ankylosing Spondylitis: New Clinical Applications in Rheumatology* Vol. 1 (eds Calabro, J. J. & Dick, C.) 23-45 (MTP, Lancaster, 1987).
- Jones, T. A. in *Computational Crystallography* (ed. Sayre, D.) 303 (Clarendon, Oxford, 1982).
- Jones, T. A. & Thirup, S. *EMBO J.* **5**, 819-822 (1986).
- Nixon, D. F. et al. *Nature* **336**, 484-487 (1988).
- Huet, S. et al. *Int. Immun.* **2**, 311-316 (1990).
- Krieger, J. I. et al. *J. Immun.* **146**, 2331-2340 (1991).

- Maryanski, J. L., Verdini, A. S., Weber, P. C., Salemme, F. R. & Corradin, G. *Cell* **60**, 63-72 (1990).
- Sette, A. et al. *Nature* **328**, 395-399 (1987).
- Sette, A. et al. *J. Immun.* **143**, 3289-3294 (1989).
- Falk, K., Rötzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290-296 (1991).
- Jardetzky, T. S. et al. *EMBO J.* **9**, 1797-1803 (1990).
- Van Bleek, G. M. & Nathenson, S. G. *Nature* **348**, 213-216 (1990).
- Falk, K., Rötzschke, O. & Rammensee, H.-G. *Nature* **348**, 248-251 (1990).
- Rötzschke, O. et al. *Nature* **348**, 252-254 (1990).
- Durbin, R. et al. *Science* **232**, 1127-1132 (1986).
- Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. *J. appl. Cryst.* **20**, 235-242 (1987).
- Crowther, R. A. in *The Molecular Replacement Method* (ed. Rossmann, M. G.) (Gordon and Breach, New York, 1972).
- Lattman, E. E. & Love, W. E. *Acta crystallogr.* **B26**, 1854-1857 (1972).
- Crowther, R. A. & Blow, D. M. *Acta crystallogr.* **23**, 544-548 (1967).
- Bricogne, G. *Acta crystallogr.* **A32**, 832-847 (1976).
- Brünger, A. *X-PLOR Manual* (Version 1.5) (Yale Univ., New Haven, 1988).
- Dornmair, K., Clark, B. R. & McConnell, H. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1335-1338 (1991).
- Elvin, J. G., Cerundolo, V., Elliott, T. J. & Townsend, A. R. M. *Eur. J. Immun.* **21**, 2025-2031 (1991).
- Cerundolo, V. et al. *Eur. J. Immun.* **21**, 2069-2075 (1991).
- Sette, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3296-3300 (1989).

ACKNOWLEDGEMENTS. We thank D. Giard (MIT Cell Culture Facility), K. Svenson and M. Frayser for large-scale culturing of LG-2 cells, and P. Klimovitsky and A. Haykov for technical assistance. This work was supported by the NSF (D.R.M.), The NIH (J.L.S. and D.R.M.), the Medical Foundation/Charles A. King Trust Fund (J.C.G.), and the Howard Hughes Medical Institute (D.C.W.).

Identification of self peptides bound to purified HLA-B27

T. S. Jardetzky*, W. S. Lane†, R. A. Robinson†, D. R. Madden*‡ & D. C. Wiley*§

* Department of Biochemistry and Molecular Biology, § Howard Hughes Medical Institute, ‡ Committee on Higher Degrees in Biophysics, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

† Harvard Microchemistry Facility, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

A pool of endogenous peptides bound to the human class I MHC molecule, HLA-B27, has been isolated. Microsequence analysis of the pool and of 11 HPLC-purified peptides provides information on the binding specificity of the HLA-B27 molecule. The peptides all seem to be nonamers, seven of which match to protein sequences in a database search. These self peptides derive from abundant cytosolic or nuclear proteins, such as histone, ribosomal proteins, and members of the 90K heat-shock protein family.

CLASS I major histocompatibility complex (MHC) molecules bind short peptides derived from intracellular proteins, forming complexes which are transported to the cell surface and accessible to immune system surveillance. In an infected cell, viral peptides complexed with class I MHC molecules are recognized by specific cytotoxic T-lymphocytes (CTL), and the infected cell is lysed¹⁻³. During their development in the thymus, CTL go through a complex process of positive and negative selection^{4,5}, generating a population of T cells poised to react to non-self peptides bound to self class I MHC molecules, while eliminating many self-reactive T cells. T-cell interaction with MHC molecules on thymic cells is important for this repertoire selection.

In an uninfected cell, class I MHC molecules are probably bound to peptides derived from self proteins. The peptide-binding sites of three human class I MHC molecules contain significant extra electron density, potentially because of bound endogenous peptides, as observed by X-ray crystallographic methods^{6-8,45}. Class I heavy- and light-chain assembly and stable cell-surface expression depends on the binding of short peptides⁹⁻¹⁵. Empty class I MHC molecules are much less stable than complexes formed with antigenic viral peptides, and short peptides can induce a conformational change in the class I heavy chain¹⁶. Finally, low relative molecular mass (M_r) material eluted from purified class I molecules could be interpreted as a complex mixture of potential self peptides^{17,18}, with peptide motifs which differed between MHC alleles (ref. 18 and G. Van Bleek and S. Nathanson, in preparation).

Here we present the sequences of 11 self peptides isolated from a pool of peptides bound to the human class I MHC molecule, HLA-B27. Seven of the peptides can be matched to proteins found in sequence databases and derive from intracellular proteins. Alignment of the self peptides with HLA-B27-restricted viral peptides reveals a sequence motif, which can be complementary to the antigen-binding site seen in the HLA-B27 crystal structure⁸.

Isolation of peptides bound to HLA-B27

Peptides were acid-eluted from purified papain-solubilized HLA-B27 (ref. 19) and recovered after ultrafiltration (Fig. 1), followed by amino-acid analysis. Amino-acid amounts 15–30-fold above background were found after acid treatment of

HLA-B27. Assuming the length of bound peptides to be 8–9 amino acids^{17,18,20}, the amount of peptidic material in the acid-eluted pool corresponds to a 50% (molar) yield, although HLA-B27 fragments also contribute to this estimate. An additional fraction, which remains to be fully characterized, was collected after reduction of the denatured HLA-B27. The combined yield from both pools corresponds to about 65% of the peptidic material expected.

The acid-eluted pool was analysed by automated Edman degradation²¹. The pool sequence did not show any dominant peptide sequence, but in cycle 2 there was 20 times more arginine than all other amino acids. Thus, arginine is probably an anchor residue¹⁸, important for peptide binding to HLA-B27.

The acid-eluted pool was fractionated by reversed-phase HPLC (Fig. 2). Single peaks from these separations were collected and microsequenced by automated Edman degradation²¹. Some of these peaks contained more than one peptide, allowing both a primary and a secondary sequence to be determined. The confidence with which a given amino-acid assignment was

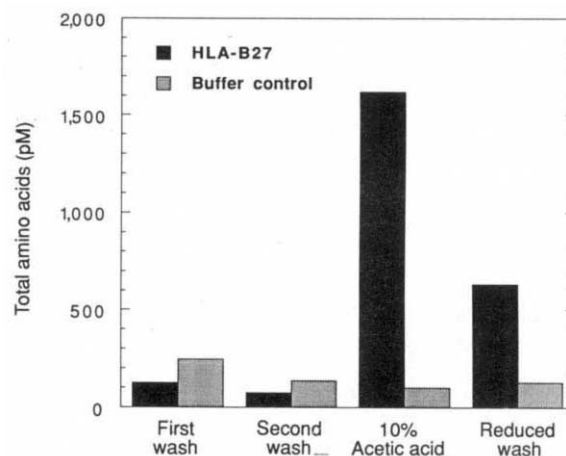


FIG. 1 Isolation of peptides bound to purified HLA-B27. HLA-B27 (B*2705) was papain-cleaved and purified as described¹⁹ from IG-2 lymphoblastoid cells. Concentrated HLA-B27 was further purified by HPLC gel filtration using a Waters SW-300 gel filtration column run in 25 mM MES pH 6.5, 150 mM NaCl, 0.1% NaN₃. About 300 µg HLA-B27 was spin-concentrated using an Amicon Centricron-10 (M_r cutoff of 10K) ultrafiltration device, to a final volume of 60 µl. Ammonium acetate (1 ml, 50 mM, pH 7.5) was added and concentrated. The flow-through provided the first wash. This procedure was repeated (second wash). Acetic acid (1 ml, 10%) was added, the sample was heated at 100 °C for 5 min and concentrated. The flow-through provided the acid-eluted low M_r pool. An equivalent volume (60 µl) of 10% ammonium hydroxide was added to the retentate followed by 1.2 µl β-mercaptoethanol. At this point a precipitate was observed. The sample was incubated for 1 h at 37 °C. Acetic acid (1 ml, 10%) was added, at which point the precipitate dissolved. The sample was spin-concentrated and the flow-through collected. The four pools were concentrated to 100 µl in a Savant Speed-Vac and lyophilized. The pools were then dissolved in 100 µl water and 5% was analysed for amino-acid content on an ABI 420A/130A derivatizer/HPLC after hydrolysis *in vacuo* with 6M HCl for 24 h.

TABLE 1 Peptide sequences

Peak	Yield (pmol)	Peptide sequence										Homologous protein		
ELUTED PEPTIDES														
2-27	16	R	R	Y	Q	K	S	T	E	L				Human histone H3, H3.3
	4	R	R	I	K	E	I	V	K	K				Human Hsp89 α
2-14	4	R	R	V	K	E	V	V	K	k				Human Hsp89 β
2-83	4	R	R	W	L	P	A	G	d	a				Human elongation factor 2
2-25a	3	R	R	S	K	E	I	T	V	R				Human ATP-dependent RNA helicase
2-42	16	G	R	I	D	K	P	I	L	K				Ribosomal protein (yeast/slime mould)
2-62b	3	F*	R	Y	N	G	L	i	H	r				Rat 60S ribosomal protein L28
1-14	1.5	K	R	F	E	G	L	T	Q	R				—
2-45	6	R	R	F	T	R	P	E	H	—				—
2-46	7	R	R	I	S	G	V	D	R	Y				—
2-62a	14	A	R	L	F	G	I	R	A	K				—
2-25b	3	[P	P	K	T	H	V	T	H	H	P	[...]		HLA-B27 (182-)
2-93	11	[G	S	H	S	M	R	Y	F	H	T	S	V	[...] HLA-B27 N terminus
VIRAL PEPTIDES														
		S †	R	Y	W	A	I	R	T	R				Influenza A nucleoprotein 383-391
		K	R	W	I	I	L	G	L	N	K	I	V	HIV GAG p24 protein 265-276
		G	R	A	F	V	T	I	G	K				HIV gp120 314-322

Isolated peptides were HPLC purified as described in Fig. 2 and microsequenced using an ABI 477A protein sequencer as in ref. 21. Yields are based on initial sequence levels. Bold, full confidence sequence data; not bold, high confidence sequence data; lower case, most significant PTH-amino acid, but less than high confidence; underlined, where peaks contain two peptides, PTH-amino-acid level was consistent with the presence of the amino acid in both sequences; brackets, peptides derived from HLA-B27, not bound in the site.

* Ser and Lys in cycle 1 and Phe in cycle 3 also consistent with sequencing data.

† Ser not present in synthetic peptides used in ref. 27.

made is divided into four different categories (Table 1). All peptide sequences were found in the 1-16 picomole range, where 30 picomole represents an occupancy of about 1% of the HLA-B27 molecules. Each of the sequences was used to search protein and nucleic-acid databases (refs 22, 23 and Sequence Analysis Software Package, Genetics Computer Group, Madison).

Peptide homology

Five of the isolated peptides match to human proteins with intracellular locations and two peptides were derived from HLA-B27 (Table 1). The primary sequence data from peak 1-8 and peak 2-27 match the human H3 and H3.3 histone proteins, whereas the secondary sequence matches a member of the human 90K heat-shock protein family, Hsp89 α . A second peptide, found in peak 2-14, matches exactly to a closely related heat-shock protein, Hsp89 β (ref. 24), and the two heat-shock peptides derive from the same region of their associated protein sequences. A peptide which matches the human elongation factor 2 protein sequence was found in peak 2-83.

In peak 2-25, two peptide sequences were observed in a 1:1 molar ratio. After identification of an acid cleavable²⁵ fragment of HLA-B27 (peak 2-25b, Table 1), the remaining peptide sequence in peak 2-25 was matched to the sequence of a human ATP-dependent RNA helicase (peak 2-25a, Table 1).

A representative large peak, eluting later in the acetonitrile gradient (Fig. 2b, peak 2-93), was sequenced and found to be a long N-terminal peptide from HLA-B27. The two HLA-B27 peptides (peak 2-25b and peak 2-93) are probably peptides generated by papain cleavage and acid treatment and not specifically bound in the antigen-binding site.

Two peptides (peaks 2-42 and 2-62) have significant sequence similarity to ribosomal proteins, whose human homologues are not in the databases. The peptide from peak 2-42 is similar to three proteins with one amino acid mismatched in each alignment: the slime mould (*Dictyostelium discoideum*) ribosomal protein L2, the yeast (*Schizosaccharomyces pombe*) ribosomal protein K37, and the yeast (*S. pombe*) ribosomal protein rpKD4. The single mismatches occurred in the same sequence position and are the conservative substitutions methionine or leucine,

instead of the experimentally determined isoleucine at position 7 (Table 1). The sequencing data suggest that the human protein will have isoleucine at this position.

In the secondary sequence data from peak 2-62, a second potential ribosomal protein peptide was found. An exact match is found with the rat ribosomal protein L28 (peak 2-62b, Table 1). This peptide sequence may be conserved in the rat and human proteins.

Four of the sequenced peptides did not match any database proteins with statistical significance. These peptides were found in peaks 1-13, 2-45, 2-46 and 2-62a and all gave high confidence sequence data. Three of the four peptides provided nine cycles of clear sequence data. The PTH amino-acid yields for peak 2-45 tapered off in the later cycles and were too low to conclude that the peptide is only eight amino acids long.

The HLA-B27 peptide-binding motif

The sequencing data from the peptide pool and the HPLC-purified peptides (Table 1) share sequence patterns, the most striking being the arginine at position 2. The interpretation that the sequence patterns reflect the specificity of the HLA-B27 peptide-binding site is supported by the observation that three viral peptides which bind to HLA-B27 (refs 26, 27 and K. Parker, *et al.*, in preparation) align well with the self-peptide sequences (Table 1). In addition, three of the self peptides (KRFEGLTQR, RRYQKSTEL, GRIDKPILK; single-letter amino-acid code) have been synthesized and stimulate the assembly of denatured HLA-B27 (ref. 28 and E. Collins, unpublished data). A tally of all the observed residues at each position of the peptides (P1-P9) is presented in Fig. 3. The number of times an amino acid is found at each peptide position is graphically represented using the structural data for HLA-B27 (ref. 8).

The most restricted peptide position is P2, with arginine found in all of the peptide sequences, followed by P1 and P9, which show a preference for positively charged amino acids. Further constraints are found at positions P3 (hydrophobic) and P6 (nonpolar or small polar residues) in the peptides. The remaining positions, P4, P5, P7 and P8 are fairly unrestricted, with both

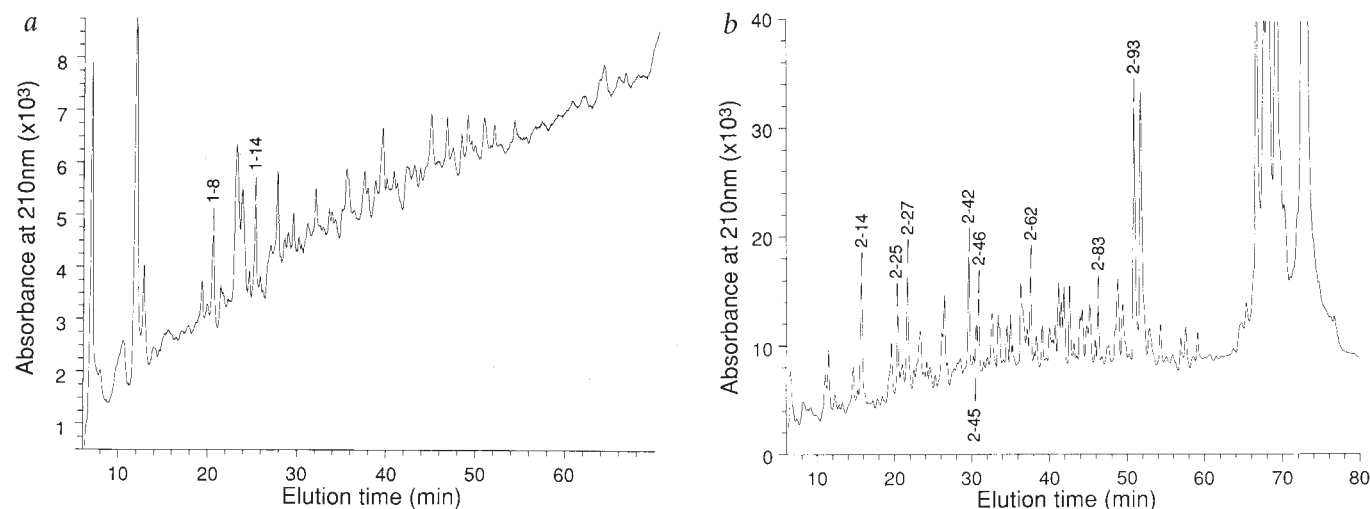


FIG. 2 Narrowbore reversed-phase HPLC separation of peptides. *a*, Five per cent of the HLA-B27 acid-eluted pool was loaded onto a 2.1 mm $\times$ 150 mm Vydac C-18 reversed-phase column. *b*, Fifty per cent of the HLA-B27 acid-eluted pool was loaded onto a 2.1 mm $\times$ 150 mm Vydac C-8 reversed-phase column. The samples were chromatographed on a Hewlett-Packard 1090

HPLC/1040 diode array detector, using a modification of a previously described gradient^{21,44}. Peaks are labelled with the prefix 1 or 2 to identify the column run and fraction number. Fourteen peaks were chosen for sequencing based on peak symmetry, resolution, ultraviolet absorbance and spectra, with 10 providing sequence data.

negatively and positively charged side chains as well as polar and nonpolar residues represented.

Conclusions

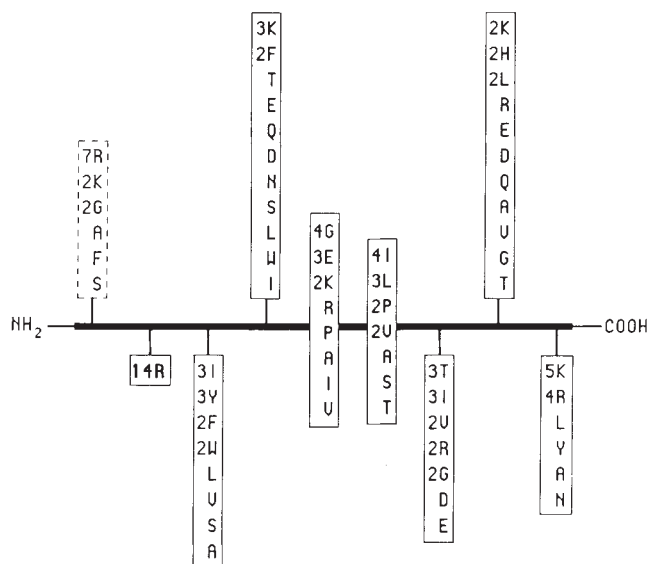
Individual peptide sequences have been determined from a pool of peptides associated with purified HLA-B27 molecules. A heterogeneous population of peptides is eluted on acid denaturation of HLA-B27 in yields consistent with stoichiometric MHC-peptide complexes. Sequence analysis of this peptide mixture reveals a predominance of arginine at cycle 2, remarkably consistent with the expected specificity constraints of the HLA-B27 peptide-binding site⁸. Narrow-bore HPLC separation allowed the isolation and microsequence analysis of 11 individual self peptides. These self peptides share a sequence motif, which can be used to align three viral peptides known to bind HLA-B27. In addition, three selected self peptides stimulate the *in vitro* assembly of HLA-B27. These data support the conclusion that the peptide sequences reported here are derived from HLA-B27

peptide complexes found in the lymphoblastoid cell line LG-2 grown in cell culture.

The compilation and alignment of 14 individual peptide sequences provides an HLA-B27 peptide-binding motif, which can be interpreted using the X-ray crystallographic structure of HLA-B27 (ref. 8). Class I MHC molecules must bind a large number of diverse peptides to guarantee T-cell mediated immunity to infection. Assuming each peptide position is completely independent, the limited motif data in Fig. 3 represent a potential for binding ($6 \times 8 \times 11 \times 8 \times 7 \times 7 \times 11 \times 6$), or 13 million, unique peptide sequences.

Seven of the self peptides are similar to intracellular proteins, with either cytosolic (ribosomal proteins, elongation factor 2 and heat-shock proteins, ref. 29) or nuclear (histone, helicase, ref. 30) locations. This intracellular localization is consistent with the observation that class I MHC molecules bind antigenic peptides from non-self proteins degraded in the cytoplasm³¹⁻³³, and that cells with a defective peptide transport mechanism

FIG. 3 The HLA-B27 peptide binding motif. The peptide sequences shown in Table 1 were used to compile a list of the amino acids found at positions P1-P9 for nonamer peptides bound to HLA-B27. Residues which occur once at a position are listed in single-letter amino-acid code, whereas multiple occurrences are indicated by number. P1-P9 are represented from left to right, with the far left equivalent to the N terminus of the peptide. The placement of the boxed peptide positions reflects information based on the analysis of the HLA-B27 structure⁸. Boxed positions which point down (P2, P3, P7, P9) are those for which observed extra electron density points down into HLA-B27 pockets. Boxed positions which point up (P4 and P8) are those which point directly out of the HLA-B27 peptide binding site. Boxed positions which are centred (P5 and P6) are those for which side chains lie across the peptide-binding site. The dashed box for peptide position P1 reflects the fact that no electron density was observed for this peptide side chain. But model building suggests that the longer side chains of arginine and lysine at P1 can point up from the binding site.



express low amounts of class I molecules on their surface^{9,34-38}. No obvious protease cleavage sites were found in common in the self-protein sequences, which could generate the observed self peptides.

Most of the identified self peptides derive from abundantly expressed proteins. Histones and ribosomal proteins are present in millions of copies in the cell^{39,40}. Elongation factor 2 is similar to Ef-Tu, one of the most abundant proteins in *Escherichia coli*⁴⁰, and the heat-shock proteins Hsp89 α and Hsp89 β are abundantly and constitutively expressed in cells cultured in serum^{24,29}. Highly expressed proteins might be expected to provide more peptide for class I MHC molecules than poorly expressed proteins and thereby be better represented in the self-peptide pool. But the number of peptide complexes formed and presented at the cell surface will probably depend on other factors as well.

Received 23 July; accepted 23 August 1991.

1. Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701-702 (1974).
2. Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457-467 (1985).
3. Townsend, A. R. M. & Bodmer, H. A. *Rev. Immun.* **7**, 601-624 (1989).
4. Strominger, J. L. *Science* **244**, 943-950 (1989).
5. Von Boehmer, H. A. *Rev. Immun.* **8**, 531-556 (1990).
6. Bjorkman, P. J. *et al. Nature* **329**, 506-512 (1987).
7. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692-696 (1989).
8. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **353**, 321-325 (1991).
9. Townsend, A. R. M. *et al. Nature* **340**, 443-448 (1989).
10. Townsend, A. R. M. *et al. Cell* **62**, 285-295 (1990).
11. Schumacher, T. N. M. *et al. Cell* **62**, 563-567 (1990).
12. Lie, W. R. *et al. Nature* **344**, 439-441 (1990).
13. Cerundolo, V. *et al. Nature* **345**, 449-452 (1990).
14. Ljunggren, H. G. & Karre, K. *Nature* **346**, 476-480 (1990).
15. Kvist, S. & Hamann, U. *Nature* **348**, 446-448 (1990).
16. Elliott, T., Cerundolo, V., Elvin, J. & Townsend, A. *Nature* **351**, 402-406 (1991).
17. Van Bleek, G. M. & Nathenson, S. G. *Nature* **348**, 213-216 (1990).
18. Falk, K., Rötzschke, O., Stevanović, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290-296 (1991).
19. Parham, P., Alpert, B. N., Orr, H. T. & Strominger, J. L. *J. biol. Chem.* **252**, 7555-7567 (1977).
20. Rötzschke, O. *et al. Nature* **348**, 252-254 (1990).
21. Lane, W. S., Galat, A., Harding, M. W. & Schreiber, S. L. *J. Prot. Chem.* **10**, 151-160 (1991).
22. Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444-2448 (1988).
23. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. molec. Biol.* **215**, 403-410 (1990).
24. Hickey, E., Brandon, S. E., Smale, G., Lloyd, D. & Weber, L. A. *Molec. cell. Biol.* **9**, 2615-2626 (1989).
25. Terhorst, C., Robb, R., Jones, G. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4002-4006 (1977).
26. Nixon, D. F. *et al. Nature* **336**, 484-487 (1988).

All the processes which alter the origins and the amounts of self peptides bound to class I MHC molecules at the cell surface could have effects on T-cell development, the allorecognition of non-self MHC (ref. 46), the induction of autoimmune disease, and the immune surveillance of infected or cancerous cells. Protein degradation rates affect viral peptide presentation to CTL (ref. 41) and the specific expression of a normal protein may affect tumour rejection⁴². The mouse homologue of the human Hsp89 α protein has also been identified as a tumour transplantation antigen⁴³.

Note added in proof: All of the large peaks to the right of peak 2-93 in Fig. 2b contain peptides longer than 12 amino acids, with sequences starting at the N-terminus of the HLA-B27 heavy chain or β_2 -microglobulin, or the acid-cleaned band of the heavy chain. $\square$

27. Huet, S. *et al. Int. Immun.* **2**, 311-316 (1990).
28. Silver, M. L., Parker, K. C. & Wiley, D. C. *Nature* **350**, 619-622 (1991).
29. Lai, B.-T., Chin, N. W., Stanek, A. E., Keh, W. & Lanks, K. W. *Molec. cell. Biol.* **4**, 2802-2810 (1984).
30. Iggo, R. D. *et al. Molec. cell. Biol.* **11**, 1326-1333 (1991).
31. Morrison, L. A., Lukacher, A. E., Braciale, V. L., Fan, D. P. & Braciale, T. J. *J. exp. Med.* **163**, 903-921 (1986).
32. Moore, M. W., Carbone, F. R. & Bevan, M. J. *Cell* **54**, 777-785 (1988).
33. Townsend, A. R. M., Bastin, J., Gould, K. & Brownlee, G. G. *Nature* **324**, 575-577 (1986).
34. Spies, T. *et al. Nature* **348**, 744-747 (1990).
35. Monaco, J. J., Cho, S. & Attaya, M. *Science* **250**, 1723-1726 (1990).
36. Trowsdale, J. *et al. Nature* **348**, 741-744 (1990).
37. Deverson, E. V. *et al. Nature* **348**, 738-741 (1990).
38. Spies, T. & DeMars, R. *Nature* **351**, 323-324 (1991).
39. Alberts, B. *et al. in Molecular Biology of the Cell* 2nd edn: 496 (Garland, New York, 1989).
40. Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A. & Weiner, A. M. *in Molecular Biology of the Gene* 4th edn: 466, 652, 677, 725 (Benjamin/Cummings, Menlo Park, 1987).
41. Townsend, A. R. M. *et al. J. exp. Med.* **168**, 1211-1224 (1988).
42. Van den Eynde, B., Lethé, B., Van Pel, A., De Plaen, E. & Boon, T. *J. exp. Med.* **173**, 1373-1384 (1991).
43. Ullrich, S. J., Robinson, E. A., Law, L. W., Willingham, M. & Appella, E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3121-3125 (1986).
44. Stone, K. L. *et al. in Techniques in Protein Chemistry* (ed. Hugli, T.) 377-391 (Academic, San Diego, 1989).
45. Saper, M. A., Bjorkman, P. J. & Wiley, D. C. *J. molec. Biol.* **219**, 277-319 (1991).
46. Matzinger, P. & Bevan, M. J. *Cell Immunol.* **29**, 1-5 (1977).

ACKNOWLEDGEMENTS. We thank J. Gorga and J. Strominger for their initial and continuing collaboration on the HLA-B27 crystal structure and A. Haykov for technical assistance. L. Stern for helpful discussions. E. Collins and K. Parker for sharing unpublished data and M. Frayser and K. Svenson for growing LG-2 cells. The work was supported by the Howard Hughes Medical Institute. T.S.J. is a postdoctoral fellow of the Cancer Research Institute.

LETTERS TO NATURE

Evidence that the compact object in SS433 is a neutron star and not a black hole

Sandro D'Odorico*, Tom Oosterloo*, Tomaž Zwitter† & Massimo Calvani‡

* European Southern Observatory, D-8046, Garching, Germany

† University of Ljubljana, Department of Physics, Ljubljana 61000, Yugoslavia

‡ Astronomical Observatory, Padua 35122, Italy

THE unusual galactic object SS433 is well known because of the periodic red and blueshifts, corresponding to velocities of 50,000 km s⁻¹, of some of its emission lines^{1,2}. It is now believed to be a binary system that emits two oppositely directed precessing jets moving with a speed of 0.26c. The jets are produced and controlled by an accretion disk, probably geometrically thick, around a compact object whose nature is still controversial. Several arguments have been advanced²¹⁻²³ suggesting that it is a black hole. Here we report spectroscopic observations of the He II line at 4,686 Å from which we deduce a new estimate of the orbital speed of the compact object. Together with the mass ratio of the binary components, derived from X-ray observations, we find that the compact object is a neutron star, not a black hole. Analysis

of the double-peaked profile of the He II line suggests that it is emitted by the accretion disk (or its corona), which is partly obscured by an opaque wind from the hot-spot region.

The amplitude K of the orbital variation of the He II line at 4,686 Å is a fundamental parameter of SS433 because, when interpreted as a tracer of the orbital mass function to be obtained:

$$F = 1.035 \times 10^{-7} K^3 \times P_{\text{orb}} (\sin i)^{-3} M_{\odot} \\ = \frac{M_x}{q(1+q)^2} = \frac{M_{\star}}{(1+q)^2} \quad (1)$$

(where K is in km s⁻¹, and the orbital period P_{orb} is in days; $M_{\star}$ is the mass of the normal star and M_x is the mass of the compact object) and essentially determines the nature of the compact object, once an estimate for the mass ratio q is provided (values of P_{orb} and the system inclination i are given in ref. 3). The need for a detailed study of the He II line at 4,686 Å has been outlined in recent reviews³⁻⁶ of this interesting system. The first determination of the Doppler modulation⁷ gave a rather high amplitude, $K = 195 \pm 19$ km s⁻¹, corresponding to $F = 10.6 M_{\odot}$. The complex and unresolved profile of their spectra made a separation of true orbital modulation from strength variation of line components difficult.

To disentangle the values of the masses of the binary components, one needs an estimate of the mass ratio q . This can be inferred from X-ray observations of the primary eclipse⁸⁻¹¹: in

TABLE 1 Sinusoidal fits to orbital variations

	K	Φ_0	γ	σ
Line centroid	112 ± 5	-0.08 ± 0.01	$+43 \pm 5$	34
Blue peak	42 ± 10	-0.14 ± 0.06	-192 ± 12	36
Red peak	133 ± 24	-0.14 ± 0.03	$+305 \pm 21$	72
Intensity ratio	0.21 ± 0.03	-0.19 ± 0.05	1.00 ± 0.02	0.05

The orbital variations are fitted to a function $(-K \times \sin 2\pi(\Phi + \Phi_0) + \gamma)$. σ denotes standard deviation of the observed and fitted values. K , γ and σ are in km s^{-1} (except for the ratios in the final line). The intensity ratio is for the red peak to the blue peak.

contrast to many contributions to the optical emission, most of the X-ray flux comes from a small region at the base of the relativistic jets²⁴. Such an 'X-ray candle' follows the orbital motion of the compact object. When it is eclipsed by the normal star (primary eclipse), the duration of the eclipse is related to the relative size of the normal star and in turn to the mass ratio q (the normal star is assumed to fill its Roche lobe). These estimates suggest^{10,12} $q \approx 0.245$ if the X-ray jets are partly obscured by the accretion disk and (during eclipses) by the normal star. Using the value of the mass function derived by Crampton and Hutchings⁷, one then gets the mass of the normal star $M_* = 16.4 M_\odot$ and the mass of the compact object $M_x =$

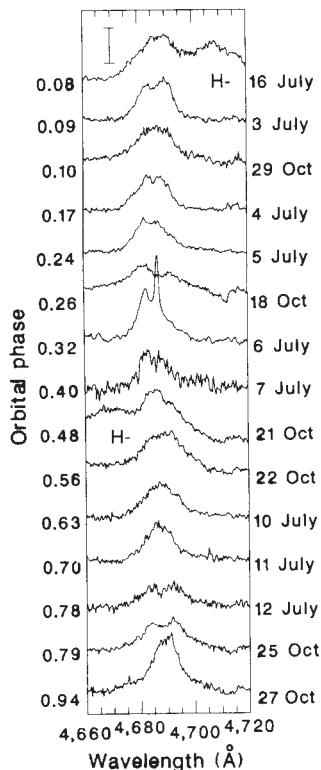


FIG. 1 Intensity tracings of the region near the 4,686 Å line of He II. Our observations were obtained with the EMMI spectrograph on the New Technology telescope in the blue channel, with a grating of 1,200 grooves per mm, working in first order (EMMI no. 3). The spectral interval covered by one exposure is 32 nm, centred on the line at 4,686 Å. The detector was a $1,024 \times 1,024$ pixel Thomson CCD, dye-coated to enhance the ultraviolet and blue sensitivity. The exposure times were between 30 and 60 min. The resolution of the spectra is ~ 0.9 Å, as measured from the width of the lamp lines used for comparison. The signal-to-noise ratios in the extracted spectra are generally $\sim 30:1$ for each wavelength bin in the line, and the wavelength calibration is accurate to ≤ 0.03 Å. Dates of observations and orbital phases are marked. Note the passages of the H_{β} -moving line. A cosmic ray is superimposed on the red peak of the spectrum taken on 6 July.

$4.0 M_\odot$. These results suggest that the compact object is a black hole. Note, however, that the mass function is proportional to K^3 , therefore its value must be precisely known if reliable values are to be obtained for the masses of the components. Indeed, Crampton and Hutchings commented that the profiles of the He II emission line may be complex, and that if this is so, the effect will be to decrease the K amplitude and the mass function.

We present here data derived from the analysis of 15 high-quality spectra obtained with the EMMI spectrograph on the European Southern Observatory's 3.5-m New Technology Telescope in July and October 1990. The instrument is described in ref. 13. Details of the observations are given in Fig. 1. The spectra cover two orbital cycles at precessional phases $\Psi \approx 0.16$ and $\Psi \approx 0.80$, implying that in both cases the angle between the jet axis and the line of sight was $\sim 70^\circ$ as derived from the standard kinematical model for the precessing jets.

The most important result presented here is the measurement of the orbital modulation of the whole line. The line is resolved (see Fig. 1) and in half of the spectra it shows a clear double-peaked profile. The rest of the spectra are generally more noisy, but we believe that the absence of double peaks is probably intrinsic to the source. Usually three gaussian curves were adequate to fit each line profile (as in ref. 14). The centre of the line is best represented by the centroid at half height of the fitted profile; higher centroids suffer from scatter because of the changing positions of both peaks, and lower heights are influenced by noise in the wings of the line and by uncertain determination of the continuum level. The binary orbit is circular¹⁵, so we used a least-square sine fit to determine the orbital modulation. The results are given in Fig. 2 and Table 1. We find that the amplitude of the Doppler modulation is $K = 112 \pm 5 \text{ km s}^{-1}$, corresponding to the mass function $F = 2.0 \pm 0.3 M_\odot$. These values are lower than those derived by Crampton and Hutchings⁷

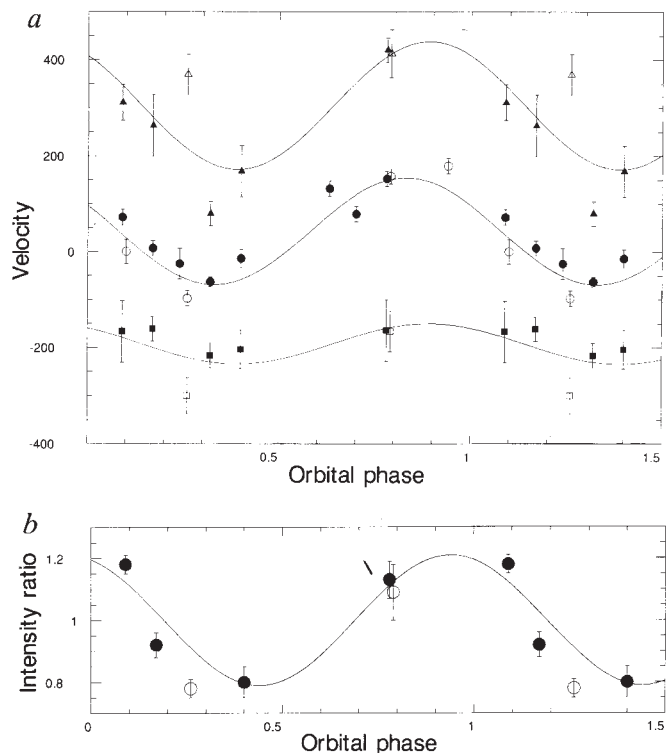


FIG. 2 *a*, Radial velocity (in km s^{-1}) of the He II line centroid at half height (circles), its red peak (triangles) and its blue peak (squares) as a function of the 13-day orbital phase²⁰ (phase zero is centred on the primary photometric eclipse). Spectra obtained in July 1990 are denoted by filled and the ones in October 1990 by open symbols. Sinusoidal fits given in Table 1 are plotted. *b*, Same as *a* but for the intensity ratio of red to blue peak.

(as they predicted, given the complex profiles); this can be explained by a precessional variation of unresolved components of the He II profile in their analysis.

The corresponding masses of the compact object and the normal star can be easily derived from equation 1: we find $M_{\star} \approx 3.2 \pm 0.4 M_{\odot}$ and $M_{\text{c}} \approx 0.8 \pm 0.1 M_{\odot}$. The above estimates for F and q , taken at face value, imply therefore that the compact object is a neutron star rather than a black hole. Given the high quality of our spectra, we are fairly confident of the value of K . A somewhat higher mass ratio could be inferred considering, as suggested in ref. 16, a moderate flattening of the secondary or a wind-type mass flow that would obscure the X-rays for a small fraction of the eclipse.

The quality of our data also allows us to study the position and strength of the line peaks. For this purpose we chose seven spectra with clear double-peaked structure. Note that the peaks are rather weak (Fig. 1) so that the line always shows a strong central component with the orbital variation described above. Sine fits to the position of the two underlying gaussians and to the intensity ratio of the two peaks are given in Table 1. We find (Fig. 2) that the red peak is stronger when the line is redshifted and the blue peak is stronger when the line is blue-shifted. The position of the blue peak seems to vary much less than the position of the red peak.

Unequal orbital modulation of the two peaks seems to support the idea of two unrelated emission regions, one approaching us and the other receding at several hundred kilometres per second. But the He II line at 4,686 Å is produced in a hot environment, so one of the peaks is likely to be emitted in the binary system region which should have small systemic velocity.

In cataclysmic variables as well as X-ray binaries, the He II emission comes from the vicinity of the compact object or, more precisely, from the corona surrounding the accretion disk. We believe that this is the case in SS433 as well. The peak separation of ~ 7.8 Å corresponds to the keplerian radius of $\approx 3 \times 10^{11}$ cm for a neutron star that is smaller than the radius of its Roche lobe.

The intensity ratio of the two peaks is such that the part of the accretion disk facing the normal star is always dimmer. This rules out the illumination from the normal star as a possible source of different peak strengths. Similarly, the eclipse of the accretion disk is far too short to cause this effect. Another possibility would be the contribution by emission from the hot spot; it is unlikely, however, that the hot spot lies on the part of the accretion disk facing away from the secondary. Finally, the emission of the He II line from the part of the corona facing the donor star might be partly absorbed by a strong wind from the hot-spot region. The existence of such outflow has been inferred from the analysis of photometric light curves¹⁷ and X-ray observations¹⁸. The rotation of the disc together with such periodic obscuration can explain both the strength and the separation modulation of the peaks. On the other hand, the high scatter of the redshifted line modulation and the large separation of the peaks on the spectrum obtained on 18 October indicate additional nonperiodic effects. Double peaks have also been observed in the Paschen stationary lines¹⁹. The splitting is smaller than observed in the He II line. This is consistent with the picture of lower-excitation lines being emitted in the outer parts of the disk (which are cooler and rotate more slowly) or in the wind from it.

In conclusion, the lower value of K measured here, together with the estimate of the mass ratio from X-ray observations of the primary eclipse, lead us to believe that the compact object in SS433 has a lower mass than previously thought. The double-peak structure observed in the He II line at 4,686 Å is here interpreted as due to emission from the corona of the accretion disk. □

3. Margon, B. *Ann. Rev. Astr. Astrophys.* **22**, 507–536 (1984).
4. Vermeulen, R. thesis, Univ. of Leiden (1989).
5. Zwitter, T., Calvani, M., Bodo, G. & Massaglia, S. *Fundamentals cosm. Phys.* **13**, 309–355 (1989).
6. Kundt, W. *Comments Astrophys.* (submitted).
7. Crampton, D. & Hutchings, J. B. *Astrophys. J.* **251**, 604–610 (1981).
8. Stewart, G. C. et al. *Mon. Not. R. astr. Soc.* **228**, 293–303 (1987).
9. Kawai, N., Matsuoka, M., Pan, H.-C. & Stewart, G. C. *Publ. astr. Soc. Japan* **41**, 491–507 (1989).
10. Zwitter, T. & Calvani, M. *Mon. Not. R. astr. Soc.* **236**, 581–588 (1989).
11. Zwitter, T. & Calvani, M. *Proc. 23rd ESLAB Symp. Two Topics in X-ray Astronomy* 701–703 (ESA SP-296, ESA, Noordwijk, The Netherlands, 1989).
12. Zwitter, T. thesis, International School for Advanced Studies, Trieste (1990).
13. Dekker, H., D'Odorico, S. & Melnick, J. *ESO Operating Manual* No. 14 (ESO, Garching, 1991).
14. Falomo, R., Boksenberg, A., Tanzi, E. G., Tarengi, M. & Treves, A. *Mon. Not. R. astr. Soc.* **224**, 323–334 (1987).
15. Collins II, G. W. & Newsom, G. H. *Astrophys. J.* **308**, 144–151 (1986).
16. Brinkmann, W., Kawai, N. & Matsuoka, M. *Astr. Astrophys.* **218**, L13–L16 (1989).
17. Zwitter, T., Calvani, M. & D'Odorico, S. *Astr. Astrophys.* (in the press).
18. Brinkmann, W., Kawai, N., Matsuoka, M. & Fink, H. H. *Astr. Astrophys.* **241**, 112–126 (1991).
19. Filippenko, A. V., Romani, R. W., Sargent, W. L. W. & Blandford, R. D. *Astr. J.* **96**, 242–250 (1988).
20. Gladyshev, S. A., Goranskii, V. P. & Cherepashchuk, A. M. *Sov. Astr.* **31**, 541–551 (1987).
21. Antokhina, E. A. & Cherepashchuk, A. M. *Sov. Astr. Lett.* **11**, 4–7 (1985).
22. Leibowitz, E. M. *Mon. Not. R. astr. Soc.* **210**, 279–301 (1984).
23. Leibowitz, E. M., Mazeh, T. & Mendelson, H. *Nature* **307**, 341–342 (1984).
24. Watson, M. G., Stewart, G. C., Brinkmann, W. & King, A. R. *Mon. Not. R. astr. Soc.* **222**, 261–271 (1986).

ACKNOWLEDGEMENTS. We would like to thank A. Čadež and R. Falomo for reading the manuscript. The observations on which this paper is based were obtained during the commissioning of EMMI. We are grateful to the EMMI project team for the excellent performance of the instrument.

Observation of muons using the polar ice cap as a Cerenkov detector

D. M. Lowder*, T. Miller*, P. B. Price*, A. Westphal*, S. W. Barwick†, F. Halzen‡ & R. Morse‡

* Department of Physics, University of California, Berkeley, California 94720, USA

† Department of Physics, University of California, Irvine, California 92717, USA

‡ Department of Physics, University of Wisconsin, Madison, Wisconsin 53706, USA

DETECTION of the small flux of extraterrestrial neutrinos expected at energies above 1 TeV, and identification of their astrophysical point sources, will require neutrino telescopes with effective areas measured in square kilometres—much larger than detectors now existing^{1–3}. Such a device can be built only by using some naturally occurring detecting medium of enormous extent: deep Antarctic ice is a strong candidate. A neutrino telescope could be constructed by drilling holes in the ice with hot water into which photomultiplier tubes could be placed to a depth of 1 km. Neutrinos would be recorded, as in underground neutrino detectors using water as the medium, by the observation of Cerenkov radiation from secondary muons. We have begun the AMANDA (Antarctic Muon and Neutrino Detector Array) project to test this idea, and here we describe a pilot experiment using photomultiplier tubes placed into Arctic ice in Greenland. Cerenkov radiation from muons was detected, and a comparison of count rate with the expected muon flux indicates that the ice is very transparent, with an absorption length greater than 18 m. Our results suggest that a full-scale Antarctic ice detector is technically quite feasible.

Neutrinos would be detected by the standard technique of observing the Cerenkov light emitted by upward-moving muons, produced by neutrinos interacting in or below the detector. At these high energies, the direction of the muon produced is within 1° of the direction of the parent neutrino, so the observation of astrophysical neutrino sources is possible in principle. Such a detector at the South Pole would have many advantages: (1) unlike existing and some proposed detectors, there would be no physical limit to its size and structure, and it could be gradually expanded (holes could be added year after year); (2) the ice would form a mechanical structure for the instrument, enabling all active electronics other than photomultiplier tubes (PMTs) to be placed directly above at the surface, putting most

Received 3 June; accepted 19 July 1991.

1. Margon, B. et al. *Astrophys. J.* **230**, L41–45 (1979).
2. Mammano, A., Ciatti, F. & Vittone, A. *Astr. Astrophys.* **85**, 14–19 (1980).

of the electronics within reach for adjustment and maintenance; (3) the lack of background light sources (for example, natural radioactivity) deep in the ice combined with low ambient temperatures imply that PMT noise rates should be very low; (4) location at the South Pole would have unique astronomical advantages—for instance, the observation of sources at constant zenith angle and the ability to observe neutrinos from a Northern Hemisphere source simultaneously with the observation of ultra-high-energy gamma rays from the same source by Northern Hemisphere air-shower arrays. For this technique to work, the ice must be very transparent (peak optical absorption length > several metres). In this initial study, conducted at the Greenland Ice Sheet Project II site (72°N, 38°W.) during 13–16 August 1990, muon rate measurements were made to assess the optical clarity of polar ice.

A string of three Thorn EMI model 9870 PMTs, each 12 cm in diameter with a hemispherical window, was lowered down a 15-cm-diameter borehole to a depth of ~217 m, ~100 m below the firn (packed snow) layer (Fig. 1). Operating close to single photoelectron sensitivity, the PMTs had noise rates between 10 and 20 kHz, higher than expected, and due to surface light leaking down the borehole. The event trigger required that all three PMTs produce a digital pulse within 30 ns of each other; the observed coincidence rate was 1.8 Hz. By inserting additional time delays to destroy the real coincidence window, the purely random coincidence rate was directly measured to be 0.2 Hz, well below the 1.8-Hz coincidence rate. From previous muon flux measurements and calculations⁴, the vertical muon flux at depth (217 m = 180 m water-equivalent) is $1 \text{ muon m}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$. The corrected event rate of 1.6 Hz, together with the expected vertical flux, indicates that each PMT was sensitive to muons

arriving within a radius of ~1 m (Fig. 2). Thus, the experiment shows that a PMT placed at a depth of 200 m in polar ice is capable of acting as a detector with an effective area >3 m². In addition, the Monte Carlo studies described next indicate that if all PMTs were operated at the single photoelectron level (not possible in Greenland because of the high noise rates), the effective area of the string would increase from ~3 to ~6 m².

Ice at this depth has a large concentration of air bubbles; the ice cores taken from this depth are milky and translucent. Davis⁵ has calculated the non-isotropic intensity distribution of light scattered by air bubbles in water, which is a good approximation to the case of air bubbles in ice. We have done Monte Carlo calculations of photon propagation in the ice, using the Davis results. In the limit where $N \geq 20$, a numerical fit to the radial distributions generated by the Monte Carlo gives $d = \lambda(1 + 3\sqrt{N})$, where d is the mean radial distance reached after N scatters, and λ is the mean free path for light scattering and $\lambda = 1/(\sigma\rho)$, where σ is the bubble cross-sectional area and ρ is the number density of bubbles. Analysis of cores taken previously from this depth in Greenland (A. J. Gow, personal communication) indicates that there are 200 bubbles cm⁻³ of mean radius 0.017 cm, giving $\lambda \sim 5.5$ cm. (For our particular experiment, $N \approx 40$, using $d = 1$ m from the rate analysis given previously and $\lambda = 5.5$ cm). These results have been incorporated into a Monte Carlo calculation to predict muon rates and pulse-height distributions for each PMT. Using the known muon flux and angular distribution, Cerenkov photons were generated from the muon paths and propagated through the bubbly ice. The program included the quantum efficiency and collection efficiency of the PMTs as measured by the manufacturer and the variation of PMT effective area with the zenith angle of an incoming photon, measured as part of the calibration of this type of PMT for the Irvine-Michigan-Brookhaven proton decay experiment⁶. The absorption length of the light in the clear ice between bubbles was scaled as a fraction of measured laboratory values⁷ (it should be pointed out that laboratory measurements of absorption include some scattering effects, and may therefore underestimate the true absorption length). Good agreement between the predicted and observed pulse-height distributions is achieved only if the peak optical absorption length of the *in situ* polar ice is greater than ~18 m, or 75% of the absorption length measured for pure laboratory ice (Fig. 3). Systematic uncertainties in the Monte Carlo calculation (for example, an

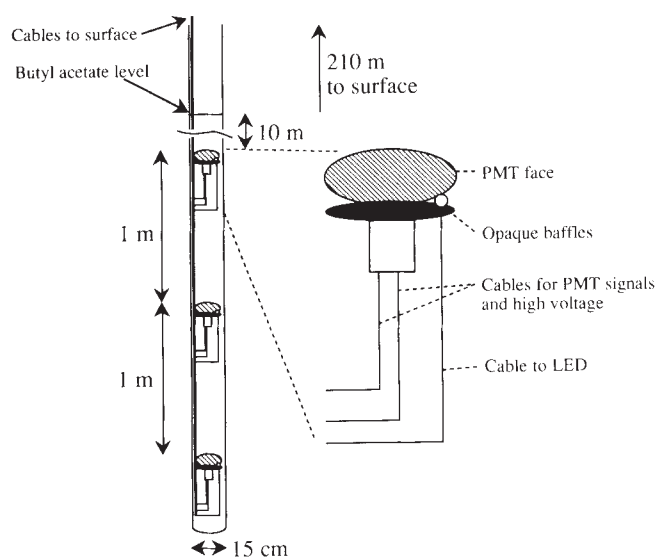


FIG. 1 PMT string used in Greenland. The three PMTs were arranged vertically, with 1-m separation between PMTs. Each PMT had a hemispherical window, and faced upwards. The hole was filled with butyl acetate, an organic liquid chosen for its low freezing point and clarity, to a depth of 11 m, for good optical contact between the PMTs and the borehole walls. The PMTs were equipped with voltage divider chains housed in pressure vessels, and high voltage was provided from the surface directly. The PMT analog output was routed through a coaxial cable (RG 180) directly to a receiver on the surface that was specifically designed by Phillips Electronics to compensate for the attenuation and dispersion introduced by the cable. Using this receiver, only a small attenuation (<25%) was observed in the amplitude of the PMT output, and the characteristic ~10 ns rise time of the PMT signal was preserved. The output signals from the receiver were routed into amplifiers, coincidence circuitry and CAMAC readout electronics. Discriminator thresholds were set as low as possible (30 mV), resulting in a sensitivity to signals greater than 1 photoelectron equivalent (p.e.) for the top PMT, 1 p.e. for the middle PMT and 2 p.e. for the bottom PMT.

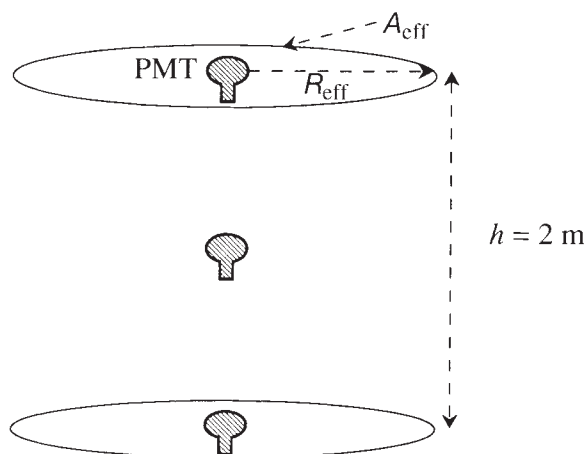


FIG. 2 The PMT string shown as a cylindrical detector, for calculation from the muon rate of the effective area of a PMT. The geometrical aperture of the instrument, in m² sr, times the muon flux ($1.0 \text{ m}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$) yields the muon rate (1.6 Hz), so the aperture is $1.6 \text{ m}^2 \text{ sr}$. Visualizing the PMT string as shown, the geometrical aperture is¹¹

$$\frac{\pi^2}{2} (2R_{\text{eff}}^2 + h^2 - \sqrt{4R_{\text{eff}}^2 h^2 + h^4})$$

This yields $R_{\text{eff}} = 1.0$ m, or $A_{\text{eff}} = 3.1 \text{ m}^2$.

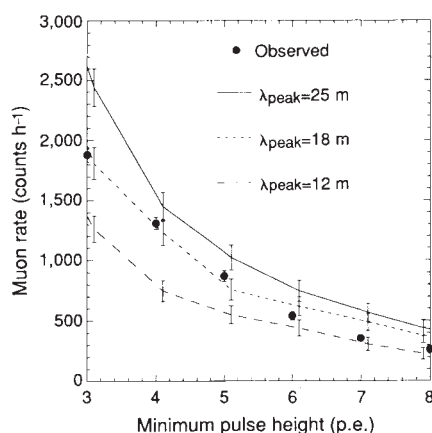


FIG. 3 Integral muon rate as a function of minimum pulse height (in p.e.) required of all three PMTs. Filled circles show experimental results, solid line shows Monte Carlo predictions if the Greenland ice has peak optical absorption length of 25 m (identical to laboratory ice). Dotted lines show Monte Carlo predictions for peak absorption lengths of 18 m and 12 m. The error bars show only statistical uncertainties in the observed rates and in the Monte Carlo rates.

observed presence of impurities in the butyl acetate used) would tend to reduce the predicted Monte Carlo rates, so our results should be regarded as a lower limit to the optical clarity of the ice. It should also be pointed out that the large absorption length implied by these results indicates that PMT size, rather than the optical quality of the ice, will be the limiting factor for a real neutrino detector: this is because for a smaller PMT, loss of light due to angular spread of the Cerenkov cone outweighs the loss of light due to optical absorption of the ice. Finally, because the ionic purity and particulate concentration of the Greenland ice are comparable to those measured for distilled water, it is not surprising that the optical properties of the two substances are also comparable. Trodahl and others⁸ have studied the diffusive transport of light through bubbly sea ice and conclude that the optical absorption length at a wavelength of 500 nm must be at least 10 m; as this ice contains brine pockets and algae, the pure ice found in the interior of Greenland or Antarctica should be much better for our purposes.

We find these results very encouraging, and are planning more extensive experiments at the South Pole during the coming austral summer (November 1991–January 1992), in particular the operation of one or two strings, each with four 20-cm-diameter PMTs at a depth of ~1 km, where the ice is expected to be bubble free⁹. Assuming that these tests are successful, we then hope to begin construction and installation of full-size strings of twenty 20-cm-diameter PMTs in the succeeding season (1992–93). In addition, as part of this year's work, we will be investigating the possibility of using our detection technique to construct muon detectors able to operate in conjunction with South Pole Air Shower Experiment, an existing air-shower array at the South Pole¹⁰. The large effective muon-sensitive area of a PMT at shallower depths suggests that good muon coverage could be achieved with only a small number of drilled holes. □

10. Smith, N. J. T. *et al.* *Nucl. Instrum. Meth. Phys. Res.* **A276**, 622–627 (1989).
11. Sullivan, J. D. *Nucl. Instrum. Meth.* **95**, 5–11 (1971).

ACKNOWLEDGEMENTS. We thank B. Koci and the entire PICO organization for the use of the borehole and for on-site assistance, E. K. Solarz and W. Williams for their help with the mechanical construction of the PMT string, J. Lynch and H. Zimmerman of the NSF, J. Learned for his sharing of DUMAND expertise, and E. Zeller of the University of Kansas for suggesting the idea of using South Pole ice in a neutrino telescope. This work was supported in part by the Division of Polar Programs of the US NSF and by the California Space Institute.

The superconducting energy gap of Rb_3C_{60}

Zhe Zhang, Chia-Chun Chen, Stephen P. Kelty, Hongjie Dai & Charles M. Lieber*

Department of Chemistry and Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

THE discovery of superconductivity in potassium-doped C_{60} (ref. 1) has been followed by an intense effort to understand the physics and chemistry of metal-doped fullerene solids^{2–13}. Experimental studies of alkali-metal-doped C_{60} have now provided insight into the structure^{7,13} and the coherence length and penetration depth⁴ of the superconducting phase. No measurements of the superconducting energy gap (Δ) have, however, been reported. The BCS theory of superconductivity¹⁵, which has been used to interpret much of this experimental work^{2,4,9–13}, predicts (in the limit of weak coupling) that the reduced energy gap $2\Delta/kT_c$ has a material-independent value of 3.53. Values in excess of 3.5 define strong coupling, and thus provide insight into the nature of the pairing mechanism. Here we describe the measurement of Δ for single-phase superconducting Rb_3C_{60} by tunnelling spectroscopy using a scanning tunnelling microscope. We obtain a value of Δ at 4.2 K of 6.6 ± 0.4 meV, corresponding to a reduced energy gap of 5.3. This is significantly larger than predicted by BCS theory, but similar in magnitude to values found for high-temperature copper oxide superconductors¹⁴. Our finding of strong coupling in Rb_3C_{60} suggests the need for caution in using standard BCS theory to interpret superconductivity in metal-doped C_{60} .

Tunnelling spectroscopy has been one of the most successful techniques used to probe conventional metal and alloy superconductors¹⁶. In particular, tunnelling spectra can provide values for both the energy gap Δ and the electron-phonon spectral function $\alpha^2F(\omega)$, where α^2 is a measure of the coupling and $F(\omega)$ is the phonon density of states. To obtain a clear value for Δ by tunnelling, it is essential to prepare a uniform insulating barrier between the superconductor and metal junction. As the C_{60} superconductors have short coherence lengths⁴ and as the samples are inhomogeneous, we expect that planar junctions might show significantly broadened gap features (even for the ideal BCS case). We have therefore used a low-temperature scanning tunnelling microscope (STM) to make point junctions with a sharpened metal tip. Point tunnelling spectroscopy has been used to measure Δ in a variety of superconducting materials, including pure metals¹⁶, organics¹⁷ and oxides¹⁸, and the values of Δ determined by this technique for conventional superconductors are comparable to those obtained with planar junctions.

The Rb_3C_{60} samples were prepared by methods described in detail elsewhere¹². Briefly, purified and dried C_{60} was reacted with RbHg in a 1:3 ratio at 200 °C for 8–72 h. After this initial reaction, the superconducting fraction determined from shielding measurements was typically >35%. The resulting powder was then ground, pressed into a pellet, and sintered at 200 °C for 3–12 h. Diamagnetic shielding measurements made on these sintered pellets show that the superconducting fraction can

Received 7 May; accepted 26 July 1991.

1. The DUMAND Collaboration Rep. No. HDC-3-88 (Hawaii DUMAND Center, 1988).
2. Barwick, S. W. *et al.* in *Proc. 3rd Int. Workshop on Neutrino Telescopes, Venice, 1991* (in the press).
3. Halzen, F., Learned, J. & Stanev, T. *AIP Conf. Proc.* **198** (eds Mullan, D. J. *et al.*) 39–51 (American Institute of Physics, New York, 1988).
4. Miyake, S. *Proc. 13th International Cosmic Ray Conference* Vol. 5, 3638–3655 (University of Denver, Colorado, 1973).
5. Davis, G. E. *J. opt. Soc. Am.* **45**, 572–581 (1955).
6. Park, H. S. thesis, Univ. Michigan (1985).
7. Warren, S. G. *Appl. Optics* **23**, 1206–1225 (1984).
8. Trodahl, H. J., Buckley, R. G. & Brown, S. *Appl. Optics* **26**, 3005–3017 (1987).
9. Gow, A. J. & Williamson, T. *Cold Regions Research and Engineering Laboratory Research Rep.* 339 (1975).

* To whom correspondence should be addressed.

approach 100% of the sample volume. After characterization of a sintered pellet by magnetic susceptibility, it was mounted on the STM sample stage in an inert atmosphere glove box, and then transferred to the STM where it was contained in a vacuum chamber. Isolation of the sample is essential in these studies, as alkali-metal-doped C_{60} decomposes on exposure to air and moisture¹⁹. The point junction was formed by mechanically moving a sharpened Pt-Ir alloy tip to the Rb_3C_{60} pellet surface and then tuning the junction resistance with a piezoceramic positioner. Current against voltage (I - V) curves were recorded digitally using custom-built electronics under computer control. The sample temperature was actively controlled for $T > 4.2$ K.

The diamagnetic shielding curve for a typical Rb_3C_{60} sintered pellet used in this study is shown in Fig. 1. The T_c of this sample is 29 K and superconducting fraction is $>60\%$; the large shielding fraction demonstrates that the sample is a bulk superconductor. After mounting a pellet in the STM, we recorded room-temperature I - V curves to verify that the sample was metallic; that is, that it had not decomposed and become semiconducting¹⁹. The STM was then cooled to 4.2 K and tunnelling spectra were recorded at different locations on the sample. An I - V curve recorded at 4.2 K on a high-quality sample is shown in Fig. 2a. This curve has a distinct zero-current regime about $V = 0$, and conductance sets in sharply at $\sim \pm 6$ mV. These features in the I - V data are the characteristic signature of the superconducting energy gap. The gap structure was observed in most of the I - V curves recorded at different locations on several high-quality samples. Occasionally we did not observe a gap; we believe, however, that such results are due to sample inhomogeneities (Z.Z. and C.M.L., unpublished data). An important observation is that the gap structure observed in Fig. 2a smears and then disappears as the sample temperature is increased above T_c (Fig. 2b, c), and furthermore, that this structure then reappears when the sample is again cooled below T_c . As the gap structure is detected reproducibly below T_c in high-quality superconducting samples, but disappears above T_c , it is clear that it can be assigned to the superconducting energy gap of Rb_3C_{60} .

We have quantitatively analysed the magnitude of Δ by calculating the conductance, dI/dV , as a function of voltage, since dI/dV is proportional to the density of states. The conductance data determined from the I - V curves show a gap in the density of states around $V = 0$ and conductance peaks at the gap edges

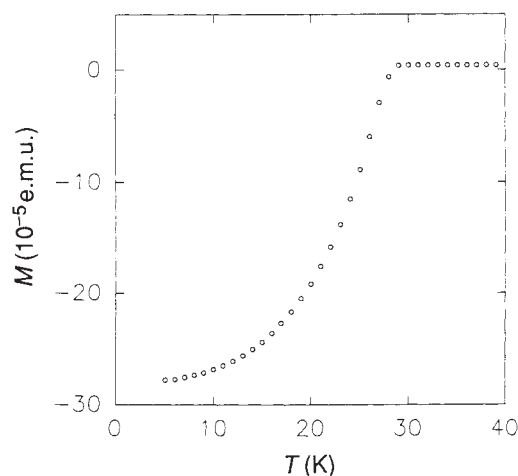


FIG. 1 Temperature dependence of the magnetization, M , obtained for a 3.1-mg Rb_3C_{60} sintered pellet. The curve was recorded by cooling in zero field to 5 K and subsequent warming in a field of 5 Oe; $T_c = 29$ K. The sample was prepared by heating a pressed pellet of Rb_3C_{60} powder for 3 h at 200 °C; the powder was first reacted at 200 °C for 72 h.

(Fig. 3). These key features agree qualitatively with our expectations based on the low-temperature BCS expression for the energy gap, $dI/dV = eV/[(eV)^2 - \Delta^2]^{1/2}$, although the quantitative fit of this expression (solid curve, Fig. 3a) to our experimental data is not ideal. We note that the value of Δ determined from this fit, 6.8 meV, is significantly larger than that predicted by BCS theory, 4.41 meV. Deviations from the ideal BCS expression are often observed in tunnelling measurements¹⁴ and could be due to inhomogeneity in our polycrystalline samples, or to broadening caused by inelastic scattering or strong coupling effects^{16,17}. In addition, it is important to recognize that it is not known whether or not these new superconducting materials should show an ideal BCS energy gap, although BCS behaviour has been implicitly assumed in many analyses of experimental data^{2,4,9-13}. To determine more accurately the value of Δ for Rb_3C_{60} , we account phenomenologically for deviations from the ideal BCS expression (Fig. 3a) by allowing for broadening as suggested by Dynes and coworkers: $dI/dV = \text{Re}[eV - i\Gamma]/[(eV - i\Gamma)^2 - \Delta^2]^{1/2}$, where Γ is the broadening term²⁰. Using this modified expression, we obtain good fits to our experimental dI/dV data with $\Delta = 6.6$ meV and $\Gamma = 0.6$ meV (Fig. 3b). It is apparent from such fits that we have observed a 'BCS-like' energy gap in Rb_3C_{60} , and that the magnitude of Δ is substantially larger than predicted by weak-coupling BCS theory.

We therefore find from a number of independent experiments that $\Delta(4.2 \text{ K}) = 6.6 \pm 0.4$ meV, that Δ becomes smaller as the temperature is increased (while remaining below T_c), and that

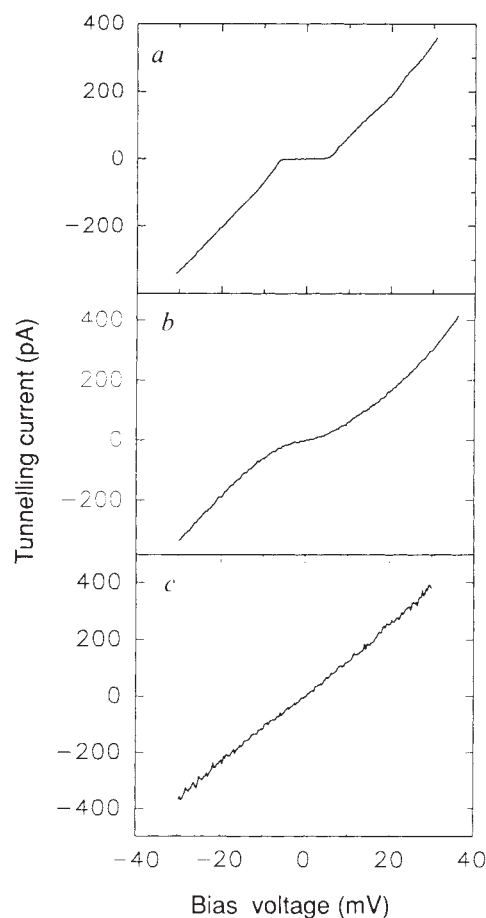


FIG. 2 Plots of I (pA) against V (mV) obtained on sintered pellet samples with a Pt-Ir alloy tip. The sample temperature in a, b and c was 4.2, 15 and 30 K, respectively. The samples for b and c were different from that in a.

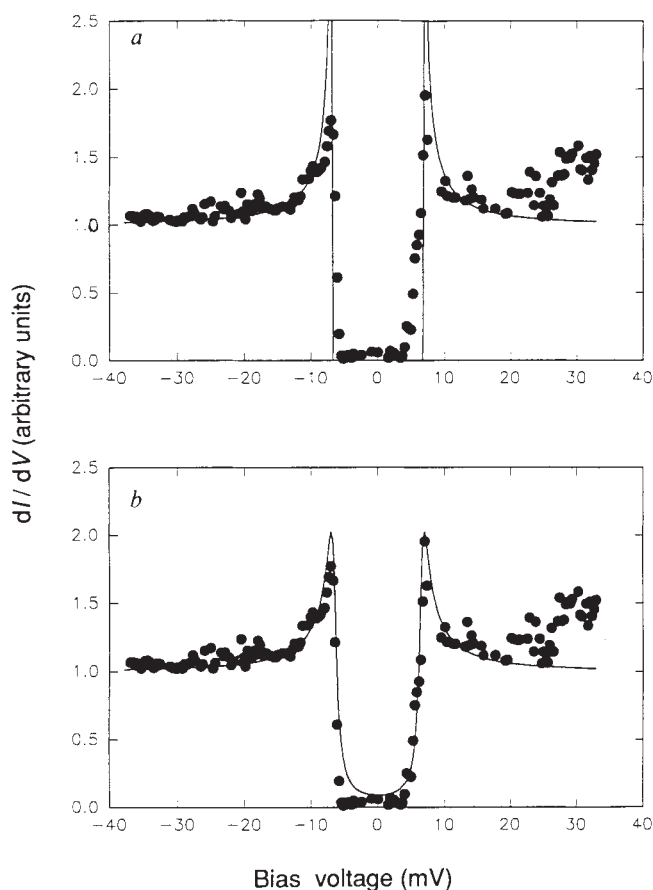


FIG. 3 Plots of dI/dV against voltage (mV) for Rb_3C_{60} at 4.2 K. The experimental data for conductance (solid circles) were calculated numerically from the $I-V$ data in Fig. 2a. *a*, The data are fitted with the expression $dI/dV = eV/[(eV)^2 - \Delta^2]^{1/2}$ (solid curve) with $\Delta = 6.8$ meV. *b*, The data are fitted with the expression $dI/dV = \text{Re}[eV - i\Gamma]/[(eV - i\Gamma)^2 - \Delta^2]^{1/2}$ which includes broadening by Γ . The values of Δ and Γ are 6.6 and 0.6 meV, respectively.

no energy gap is observed for $T \geq 30$ K. At present we do not have sufficient data to determine whether $\Delta(T)$ for Rb_3C_{60} has a BCS-like temperature dependence, although it is clear that the energy gap does not persist above T_c . We believe that our most important finding is that the value of the low-temperature reduced energy gap for Rb_3C_{60} , $2\Delta/k_B T_c = 5.3$, is significantly larger than the weak-coupling prediction of 3.53. Our data thus show that the coupling, presumably to phonons, is fairly strong in the alkali-metal-doped C_{60} superconductors. In previous studies of strong-coupling superconductors it has been noted that low-frequency phonon modes affect Δ more than high-frequency vibrations^{17,21}. Hence, it is interesting to speculate whether, in this molecular-based superconductor, C_{60} - C_{60} intermolecular modes give rise to the strong coupling determined from our tunnelling studies.

Most of the experimental data for alkali-metal-doped C_{60} superconductors have been interpreted in terms of BCS theory with variations in T_c determined by the density of electronic states at the Fermi level^{2,4,9,13}. It is apparent that these analyses should at least be re-scaled by the value of Δ measured experimentally. In addition, if T_c is indeed determined solely by the density of states at the Fermi level, we expect that the value of Δ found here for Rb_3C_{60} will be representative of C_{60} superconductors in general. It will therefore be interesting in future to investigate both the energy gap in K_3C_{60} and the electron-phonon spectral function for the alkali-metal-doped C_{60} superconductor. □

Received 27 August; accepted 4 September 1991.

1. Hebard, A. F. *et al.* *Nature* **350**, 600-601 (1991).
2. Rosseinsky, M. J. *et al.* *Phys. Rev. Lett.* **66**, 2830-2832 (1991).
3. Holczer, K. *et al.* *Science* **252**, 1154-1157 (1991).
4. Holczer, K. *et al.* *Phys. Rev. Lett.* **67**, 271-274 (1991).
5. Zhou, O. *et al.* *Nature* **351**, 462-464 (1991).
6. Benning, P. J., Martins, J. L., Weaver, J. H., Chibante, L. P. F. & Smalley, R. E. *Science* **252**, 1417-1421 (1991).
7. Stephens, P. W. *et al.* *Nature* **351**, 632-634 (1991).
8. Schirber, J. E. *et al.* *Physica C* **178**, 137-140 (1991).
9. Sparr, G. *et al.* *Science* **252**, 1829-1831 (1991).
10. Tanigaki, K. *et al.* *Nature* **352**, 222-223 (1991).
11. Kelty, S. P., Chen, C. C. & Lieber, C. M. *Nature* **352**, 223-225 (1991).
12. Chen, C. C., Kelty, S. P. & Lieber, C. M. *Science* **253**, 886-888 (1991).
13. Fleming, R. M. *et al.* *Nature* **352**, 787-788 (1991).
14. Kirtley, J. R. *Int. J. mod. Phys.* **4**, 201-237 (1990).
15. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. *Phys. Rev.* **108**, 1175-1204 (1957).
16. Wolf, E. L. *Principles of Tunnelling Spectroscopy* (Oxford University Press, New York, 1989).
17. Hawley, M. E. *et al.* *Phys. Rev. Lett.* **57**, 629-632 (1986).
18. Huang, Q. *et al.* *Nature* **347**, 369-372 (1990).
19. Haddon, R. C. *et al.* *Nature* **350**, 320-321 (1991).
20. Dynes, R. C., Narayanamurti, V. & Garno, J. P. *Phys. Rev. Lett.* **41**, 1509-1512 (1978).
21. Mitrovic, B., Leavens, C. R. & Carbotte, J. P. *Phys. Rev. B* **21**, 5048-5054 (1980).

ACKNOWLEDGEMENTS. We thank D. W. Murphy (AT&T) for communication of results before publication. C.M.L. acknowledges support of this work by the David and Lucile Packard, National Science, A. P. Sloan, and Camille and Henry Dreyfus Foundations.

Visible light emission due to quantum size effects in highly porous crystalline silicon

A. G. Cullis & L. T. Canham

DRA Electronics Division, Royal Signals & Radar Establishment, St Andrews Road, Malvern, Worcester WR14 3PS, UK

LIGHT-emitting devices based on silicon would find many applications in both VLSI and display technologies, but silicon normally emits only extremely weak infrared photoluminescence because of its relatively small and indirect band gap¹. The recent demonstration of very efficient and multicolour (red, orange, yellow and green) visible light emission from highly porous, electrochemically etched silicon^{2,3} has therefore generated much interest. On the basis of strong but indirect evidence, this phenomenon was initially attributed to quantum size effects within crystalline material², but this interpretation has subsequently been extensively debated. Here we report results from a transmission electron microscopy study which reveals the structure of the porous layers that emit red light under photoexcitation. Our results constitute direct evidence that highly porous silicon contains quantum-size crystalline structures responsible for the visible emission. We show that arrays of linear quantum wires are present and obtain images of individual quantum wires of width < 3 nm.

The current interest in fabricating low-dimensional semiconductor structures relies upon quantum size effects which can greatly alter the properties of bulk semiconductors and have, thereby, given rise to a new generation of electronic devices. It has recently been suggested^{2,4} that certain types of porous silicon can show two-dimensional carrier confinement and constitute a crystalline quantum wire array. Many workers, however, have interpreted data on porous silicon in terms of a crystallinity that deteriorates with increasing porosity and decreasing dimensionality⁵. There has even been speculation that porous silicon contains an amorphous phase in its as-anodized state⁶⁻⁹. It is, therefore, important to distinguish between these possibilities and to determine the origin of the light emission.

Porous silicon is affected by atmospheric impregnation at room temperature¹⁰, and all characterization techniques need to take this into account. Highly porous silicon is also fragile and very chemically reactive (L.T.C., unpublished data). Previous transmission electron microscope (TEM) studies¹¹⁻¹⁵ have successfully revealed the morphology of low- to medium-porosity layers in specimens prepared by sequential mechanical

polishing and ion-beam milling. The former process generally requires that the sample be waxed to a plate and treated with solvents at various times. Furthermore, the ion-milling process amorphizes the bombarded surface layers of the material to a depth of typically $\sim 3\text{--}5$ nm. As we needed to examine structures whose dimensions are a few nanometres, we considered both of these processes to be potentially harmful for the structure. Our work avoids these problems and provides evidence that visible light emission originates from quantum-size crystalline structures in the highly porous material.

The porous silicon layers studied here were all produced by electrochemical etching in '20% ethanoic HF' as described in more detail elsewhere¹⁰. One layer was then leached in an ethanol-HF solution to increase the porosity further. Wafers were dried in ambient air and cleaved for immediate photoluminescence (PL) and TEM analysis. Room-temperature PL spectra were recorded using laser illumination at 30° from the sample normal in ambient air. Front-face emission from layers was dispersed with a $\frac{3}{4}$ -m Spex monochromator and detected with a cooled germanium photodiode. Electron-transparent samples for the TEM were prepared by direct cleavage from the porous wafer surface. A sharp scalpel blade was drawn carefully across the surface, causing the porous layer to fracture and yield fragments of many sizes ejected in the direction of blade motion. The largest fragments were allowed to fall into a 400-mesh TEM specimen grid placed so that most fine, unwanted fragments overshot its position. After collection of sufficient material, gentle vibration of the grid removed any very loose fragments, whereupon the grid was inserted into the TEM for study, in general, within 1 hour of the initial fabrication of the

layer. This preparation procedure gave small flakes of material, often oriented along the (001) wafer plane, with minimal oxidation¹⁰ and minor atomic-scale damage. Examinations were carried out in a JEM 4000EX microscope using 400-keV electrons and beam fluxes as low as possible to minimize potential specimen damage.

Here, we have investigated samples of $\sim 80\%$ porosity with very weak (sample 1), medium (2) and high (3 and 4) efficiency of light emission. Samples 1–4 were fabricated by anodization of, respectively, a heavily boron-doped p-type wafer of resistivity $0.005\ \Omega\text{ cm}$ at 100 mA cm^{-2} for 1 min with subsequent leaching for 75 min; a heavily doped p-type wafer of $0.02\ \Omega\text{ cm}$ at 100 mA cm^{-2} for 1 min; a heavily arsenic-doped n-type wafer of $0.012\ \Omega\text{ cm}$ at 1.5 mA cm^{-2} for 2 h and a lightly doped p-type wafer of $40\ \Omega\text{ cm}$ at 20 mA cm^{-2} for 5 min. Figure 1 shows the room-temperature PL from samples 2–4 after ~ 1 h of air exposure. All three layers emit red light visible to the naked eye under unfocused 1 mW blue excitation. Sample 1 shows negligible visible emission: its (001) electron diffraction pattern is presented in Fig. 2a and is closely similar to that characteristic of perfect bulk material. The PL from sample 2, shown in Fig. 1, peaks at 1.53 eV with a full width at half maximum (FWHM) of 319 meV. The corresponding (001) diffraction pattern of Fig. 2b is once again similar to that of perfect bulk silicon, but there is now evidence for weak arc-shaped streaks through the main spots. Sample 3 shows much stronger luminescence at energies above that of the bulk silicon band-gap, peaking at 1.45 eV with a FWHM of 325 meV. In this case, the (001) diffraction pattern given by the layer (Fig. 2c) shows more streaking of the spots, together with greater diffuse scattering around the main beam and a very weak, narrow but diffuse ring in the position corresponding to 111 diffraction. Finally, sample 4 also emits strong red radiation at room temperature, peaking at 1.52 eV (Fig. 1) and with a significantly narrower FWHM of 244 meV. The (001) diffraction pattern from this layer shows significant electron scattering modifications. In particular, the streaking/arc-ing of the diffraction spots is very pronounced, the ring in the 111 position (marked Z in Fig. 2d) is considerably stronger and weak streaked spots corresponding to other normally unexcited Bragg reflections (such as those of type 311, marked Y) now appear.

From these observations, we note that there is a general trend, for samples 1–4, towards greater streaking and diffuse scattering

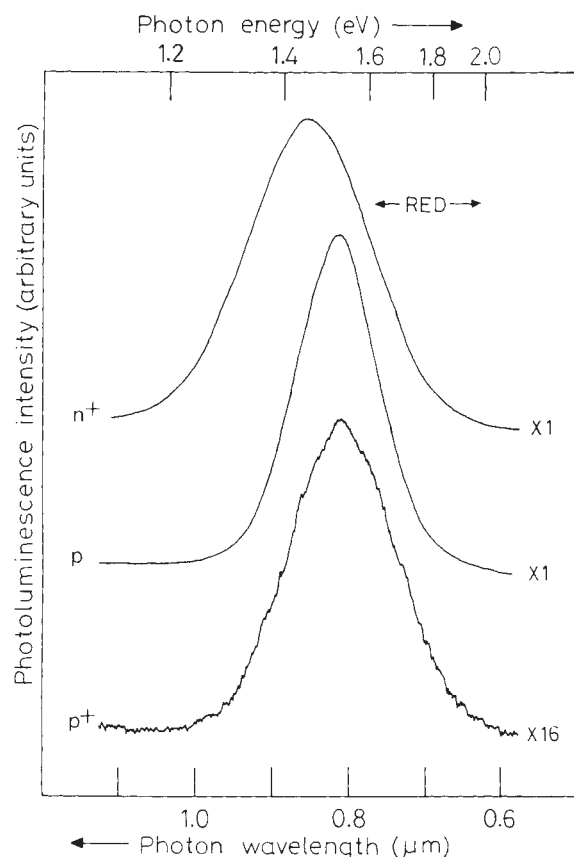


FIG. 1 Room-temperature photoluminescence from three red-emitting mesoporous layers of $78 \pm 2\%$ porosity irradiated by unfocused blue light (1 mW, 488 nm) at room temperature. Lower spectrum, sample 2, 4.2- μm layer in heavily boron-doped p-type substrate; middle spectrum, sample 4, 4.5- μm layer in lightly-doped p-type substrate; upper spectrum, sample 3, 11.6- μm layer in heavily arsenic-doped n-type substrate. Note the different scales of the spectra (right-hand side).

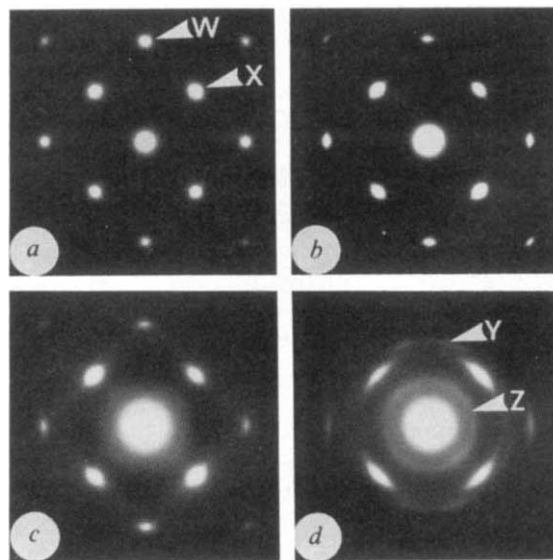


FIG. 2 Transmission electron diffraction patterns given by thin, highly porous silicon layer samples in (001) orientation. a, Sample 1 (W indicates a 400-type spot and X indicates a 220-type spot); b, sample 2; c, sample 3; d, sample 4 (Y indicates a 311-type ring and Z indicates a 111-type ring).

in diffraction patterns. There is no direct correlation between these features and radiative efficiency, however, as samples 3 and 4 produce comparable integrated output (Fig. 1) but significantly different diffraction patterns (Fig. 2). The arc-like features in the diffraction patterns would suggest that a component of misoriented material is present in some samples. Measurement of the width of the inner ring at the 111 position (Fig. 2*d*) indicates that this material has a crystallite size in a range around ~ 3 nm. At least four other factors are likely to

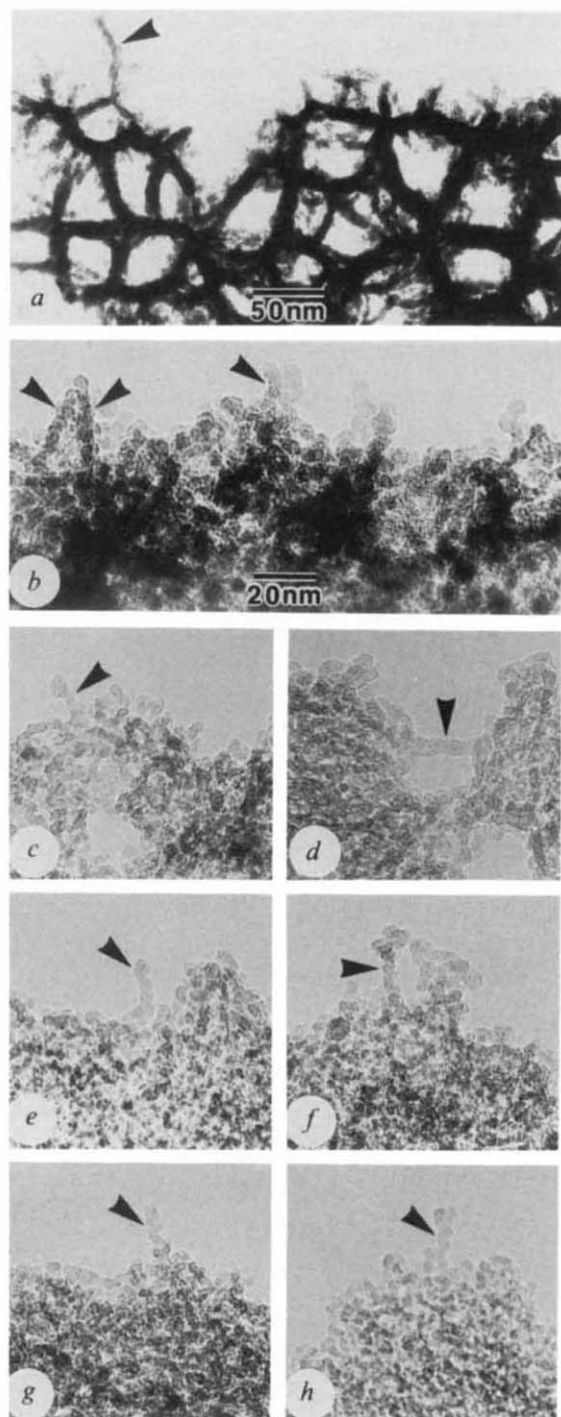


FIG. 3 Transmission electron images (bright-field with slight under-focus) of thin, porous silicon layers showing irregular matrix structures and narrow silicon columns (arrowed) protruding beyond the matrix edge. *a*, Sample 1; *b* and *c*, sample 2; *d*–*h*, sample 4. The magnification of *c*–*h* is the same as that of *b*.

give rise to additional diffraction features. The huge internal surface area of the crystalline structure ensures that a significant proportion of the silicon atoms do not lie on perfect lattice sites, and this, together with new vibrational modes of the structure, would be likely to increase diffuse scattering. Static distortions of the skeleton would probably increase diffraction angle ranges. Furthermore, from the uncertainty principle we can expect the very small dimensions of much of the crystalline silicon to give rise to at least a little broadening and streaking of spots.

Figure 3 shows the dimensions and morphology of typical, (001)-oriented crystalline silicon structures present within the highly porous layers. Sample 1 was fabricated to demonstrate the open morphology of this type of material. Figure 3*a* reveals an array of holes 25–50 nm across, often having angular or square shapes. The remaining silicon skeletal framework is single crystal with interconnecting columns that mostly have cross-sectional diameters of >10 nm, but there are some smaller interconnections (for example, the connection arrowed has a diameter of ~ 6 nm). Few columns have diameters much below this size, in accord with the very low radiative efficiency of the layer. In contrast, sample 2, despite being of similar porosity, has much smaller interstructural spacing, and overlapping features therefore obscure the details of its highly interconnected and irregular, 'coral-like' microstructure. But as shown in Fig. 3*b* and *c*, protuberances beyond the sample edge reveal that individual columns exist throughout the structure. The columns indicated in the figures have cross-sectional diameters of significantly less than 5 nm. This is direct evidence for a crystalline silicon quantum wire array. Figure 3*d*–*h* provides examples of similar isolated structures with widths down to <3 nm from sample 4. In all cases, in addition to column branching and changes in direction, it is evident that column diameters often undulate, the possible significance of which will be outlined below.

When considering the origin of the luminescence, we should note that sample 1, with a huge internal surface area, produces negligible visible emission, which implies that surface processes are not primarily linked with the phenomenon. In addition, whereas the emission of visible light from highly porous silicon has been associated previously with the existence of amorphous material, our results show that although, under some circumstances, microcrystallites are present, negligible amorphous silicon is found. Therefore, this work strongly indicates that the observed emission is due to recombination of quantum-confined carriers within crystalline silicon. From particle-in-a-box calculations of crystalline band-gap enlargement due to one-, two- and three-dimensional carrier confinement in plates, wires and dots, it is predicted that visible emission might be expected from structures of ≤ 2 nm, ≤ 3.5 nm and ≤ 5 nm dimensions, respectively (D. Whittaker and L.T.C., unpublished results). From Fig. 3, it is evident that quantum wires of the required sizes are present throughout layers that give visible luminescence. There is no direct evidence for a silicon skeleton of overall plate-like morphology. Many sections of the columnar skeleton are, however, seen to undulate in width, the extent of the variation depending on anodization conditions. Indeed, the increase in band gap associated with microscopic constrictions of quantum wires is likely to enhance carrier localization which, in turn, could enhance visible emission efficiency by retarding carrier leakage into other inactive parts of the structure. In the limit of aggressive anodization and/or leaching, sections of columns may become detached, to give the misoriented material noted from diffraction patterns. Such crystallites may give some component of dot luminescence which varies from sample to sample. It is clear from Figs 1 and 2, however, that there is no direct correlation between the presence of this material and efficiency of luminescence. We conclude that undulating crystalline quantum wires, of the type observed in our TEM images, are the structures primarily responsible for the efficient luminescence from freshly etched highly porous silicon. □

Received 8 July; accepted 6 August 1991.

- Haynes, J. R. & Westphal, W. C. *Phys. Rev.* **101**, 1676-1678 (1956).
- Canham, L. T. *Appl. Phys. Lett.* **57**, 1046-1048 (1990).
- Canham, L. T., Marsh, K. J. & Brumhead, D. *Electron. Times* **590**, 1 (1991).
- Lehmann, V. & Gösele, U. *Appl. Phys. Lett.* **58**, 856-858 (1991).
- Bomchil, G., Halimaoui, A. & Herino, R. *Micro. Eng.* **8**, 293-310 (1988).
- Turner, D. R. *J. Electrochem. Soc.* **105**, 402-408 (1958).
- Arita, Y. *J. Cryst. Growth* **45**, 383-392 (1978).
- Pickering, C., Beale, M. I. J., Robbins, D. J., Pearson, P. J. & Greef, R. *J. Phys. C* **17**, 6535-6552 (1984).
- Goodes, S. R., Jenkins, T. E., Beale, M. I. J., Benjamin, J. D. & Pickering, C. *Semicond. Sci. Technol.* **3**, 483-487 (1988).
- Canham, L. T., Houlton, M. R., Leong, W. Y., Pickering, C. & Keen, J. M. *J. appl. Phys.* **70**, 422-431 (1991).
- Phillip, F., Urban, K. & Wilkens, M. *Ultramicroscopy* **13**, 379-386 (1984).
- Beale, M. I. J., Benjamin, J. D., Uren, M. J., Chew, N. G. & Cullis, A. G. *J. Cryst. Growth* **73**, 622-636 (1985).
- Chuang, S. F., Collins, S. D. & Smith, R. L. *Appl. Phys. Lett.* **55**, 675-677 (1989).
- Sugayama, H. & Nittano, O. *J. Cryst. Growth* **103**, 156-163 (1990).
- Kaushik, V. S., Datye, A. K., Tsao, S. S., Gullinger, T. R. & Kelly, M. J. *Mater. Lett.* **11**, 109-114 (1991).

ACKNOWLEDGEMENTS. We thank D. Brumhead, P. W. Smith and K. J. Marsh for assistance in sample preparation.

A fibre-optic chemical sensor with discrete sensing sites

Steven M. Barnard & David R. Walt*

Max Tishler Laboratory for Organic Chemistry, Department of Chemistry, Tufts University, Medford, Massachusetts 02155, USA

A LONG-standing goal of clinical medicine has been the continuous, *in situ* measurement of solute concentrations (such as pH, $p\text{CO}_2$ and $p\text{O}_2$), particularly in blood. Blood monitoring is accomplished at present by analysing discrete samples at a centralized, remote clinical laboratory, causing delays between sampling and analysis. Most multi-analyte sensors developed for 'bedside' *in situ* monitoring¹ consist of several sensors fabricated into a sensor array or bundle^{2,3}. This approach is not ideal where sensor size is an important factor. Here we describe a technique that enables the development of a compact, multi-analyte fibre-optic chemical sensor. The technique is based on localized photopolymerization of appropriate dye indicators on the face of an imaging fibre. The sensing sites fluoresce to a degree controlled by analyte concentration, intensities being monitored simultaneously with an enhanced charge-coupled device (CCD) video camera. We present results from a sensor that contains three individual pH-sensitive areas, and indicate how multi-analyte sensitivity may be achieved.

The entry of fibre-optic chemical sensors into the field of analytical chemistry has provided a new approach for solving many monitoring problems⁴⁻⁷. Optical sensors are prepared by immobilizing indicators that change their optical properties on interacting with analytes. The advantages of optical sensors over their electrochemical counterparts include freedom from electromagnetic interference, lack of the need for direct electrical connections to the solution being analysed or for a reference sensor, and the potential for transmitting a higher density of information using multiwavelength transmission. This latter advantage has not been exploited fully in sensors described previously.

The useful wavelength range of optical sensors is 300-700 nm, being determined both by the fibre's higher attenuation outside these wavelengths and the lack of selective dyes above the near-infrared region. Thus, indicators useful for sensor preparation are restricted to those possessing strong absorption within this 400-nm wavelength range. These indicators are normally organic dyes that have very broad absorption and emission spectra, vitiating the use of dye combinations to measure two or more analytes with a single optical fibre because of spectral overlap. The number of parameters that may be measured simultaneously is therefore limited by the number of possible dye combinations that maintain adequate resolution at the analyte-sensitive wavelengths. Wolfbeis *et al.*⁸ have developed a multi-analyte sensor that detects both $p\text{O}_2$ and $p\text{CO}_2$ using analyte-

sensitive layers with two spectrally resolvable dyes immobilized on a single fibre tip. This technique cannot be generalized, however, because it suffers from the above restrictions. The development of multi-analyte optical sensors is limited severely if one relies on only spectral discrimination (absorption and emission characteristics). Our approach adds the dimension of spatial discrimination.

We present a technique that enables sensing regions to be placed in precise locations on the distal face of an optical fibre, the position and spectra of each being monitored simultaneously. These measurements are achieved by combining the distinct optical pathways of imaging fibres⁹ with the spatial discrimination of CCD video cameras^{10,11}. Our imaging fibres are comprised of several thousand individual fibres melted and drawn together, producing a solid single fibre that has thousands of discrete light-transmitting elements. A CCD is an integrated-circuit device which uses a two-dimensional array of photo-sensitive elements to convert incident light into a proportional electrical charge. The intensity and spatial position of features in an image projected onto the array can be measured simultaneously, and by coupling the CCD camera to appropriate optics and a properly configured computer, the images can be viewed on screen and stored on disk.

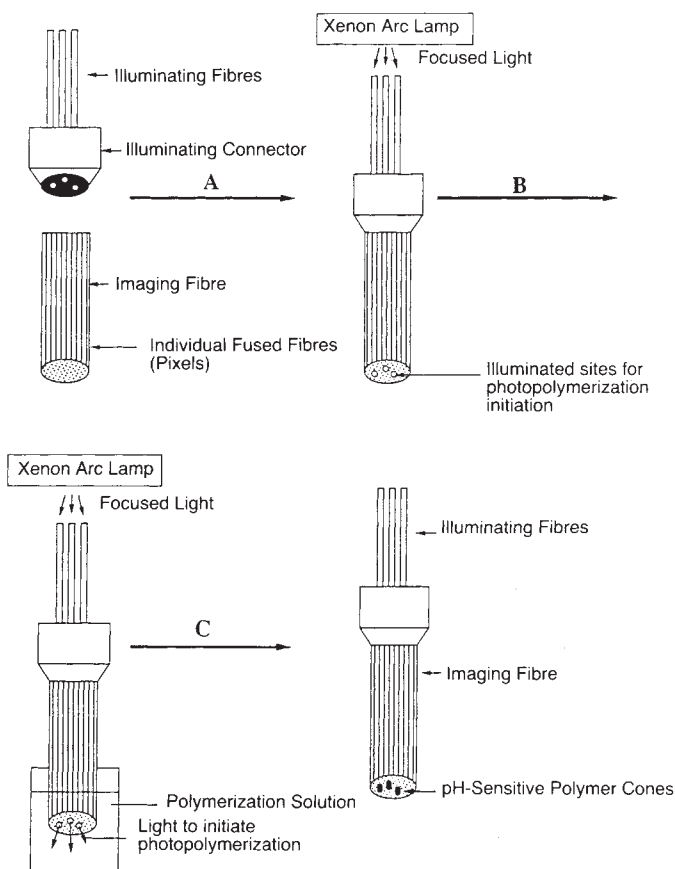


FIG. 1 Method of sensor fabrication. A 3.0-cm length of imaging fibre was silanized in (α -methylacryloxypropyl)-trimethyloxysilane¹². The imaging fibre was connected to the fibre connector, coupling the proximal end of the imaging fibre to the distal end of the illumination fibres. The illumination fibres direct light to the proximal end of a discrete number of individual pixels that comprise the imaging fibre and determine the areas of the imaging fibre's distal face that act as initiation sites for photopolymerization of the dye/polymer solution. The illumination fibres are 125 μm in diameter. The imaging fibre is suspended in a solution of acryloyl fluorescein (0.0021 M), 2-hydroxyethyl methacrylate (2.5 M), methylene bisacrylamide (0.022 M) and benzoin ethyl ether (0.12 M). The light from a xenon arc lamp is focused into the illumination fibres to initiate photopolymerization. No polymerization occurs in the bulk solution.

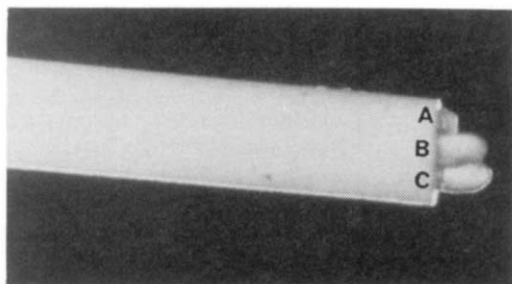


FIG. 2 Side-view photograph of a sensor with three pH-sensitive cones photopolymerized on the distal face.

To implement this strategy, it became essential to develop a technique that would place distinct sensing sites at any desired location on the distal face of an imaging fibre. This placement was accomplished by modifying our technique of growing dye-containing polymers directly on the fibre tip¹². We adapted the procedures to photoinitiation rather than thermal initiation, giving us the ability to direct the polymerization to the illuminated site only. The discrete optical pathways in the imaging fibre allow us to deliver light to a specific area and thus to induce polymerization locally.

Figure 1 shows the basic sensor fabrication scheme. Three single optical fibres (the illuminating fibres) were cemented in a connector that allowed coupling to the imaging fibre. The proximal ends of the illuminating fibres were attached to a light source and their distal ends abutted against the imaging fibre. In this manner, three distinct regions of the imaging fibre's distal face can be illuminated in direct spatial correspondence to the placement of the single optical fibres. The imaging fibre was immersed in a solution of monomers and an appropriate initiator, and photopolymerization initiated at the illuminated sites to create analyte-sensitive polymer cones on the fibre's distal face. The images transmitted to the fibre's proximal end are monitored using the CCD camera, spectral changes in the individual cones being fully resolvable spatially.

The sensors were tested on a Quantex QX7-QFM Quantitative Fluorescence Microscopy System (Sunnyvale, California) composed of an image-processing computer, monitor, RGB video monitor, intensified CCD camera, microscope and video printer. The imaging fibre containing the polymer cones was attached to the microscope stage. The fibre's distal end was submerged in a flow cell for testing of different buffer solutions. The flow cell was equipped with a 600- μm optical fibre which conveyed

light from a 75-W xenon arc lamp to the distal face of the imaging fibre. Some of the light emitted from the fluorescing cones travels up the fibre to the proximal end. The fluorescence image is focused onto the CCD through the microscope objectives and recorded by the computer.

The first sensors that we have developed contain three pH-sensitive cones fabricated by photopolymerization of acryloylfluorescein and 2-hydroxyethyl methacrylate¹³. The imaging fibre was 400 μm in diameter and comprised 1,500 individual 10- μm optical fibres (Optilec, Southbridge, Massachusetts). Fluorescein is a pH-sensitive indicator ($\text{p}K_a = 6.7$), its fluorescence increasing in basic solutions because of an increase in the quantum yield of its dianion¹⁴. A side view of the sensor photographed under a light microscope is shown in Fig. 2. The approximate dimensions of the cones (diameter by height) are: cone A, 125 μm by 70 μm ; cone B, 125 μm by 123 μm ; cone C, 125 μm by 140 μm . The cone height is determined by the photo-initiation times, which were respectively 1.0 min, 2.0 min and 3.5 min. Shorter exposure times resulted in correspondingly smaller cones.

Ten sensors were tested by immersing them in phosphate buffer solutions and recording the fluorescence image with the CCD camera. Figure 3 is a computer-generated three-dimensional plot of the transmitted fluorescence image of the fibre's distal face. The peak heights reflect the response to pH of the individual cones. The largest changes occur between pH 6.2 and 7.4, close to the $\text{p}K_a$ of fluorescein. The differences in intensity between the different cones at the same pH are due to the different amounts of fluorescein immobilized within each.

The images make clear the precise control that is possible over the location of the sensing regions. The lack of stray fluorescence on the distal face indicates the high degree of specificity that is achieved during photopolymerization. Moreover, these images show that the diameter of the illuminating fibre determines the diameter of the sensing cone, so that their number and size can be controlled. In principle, this number is limited only by the number of individual elements in the imaging fibre ($\sim 1,500$) and by the wavelength of illuminating light. Thus, micrometre-sized sensing regions should be possible using this method.

Our technique should enable the preparation of discrete sites sensitive to different analytes, thus solving many of the problems associated with designing multi-analyte optical sensors. The technique circumvents the limited wavelength range of optical fibres, overlap problems from the broad absorption and emission spectra of organic dyes, and the need to use individual fibres for each analyte. We have now prepared a dual pH/ pCO_2 sensor in which one of two pH-sensitive cones was converted to a pCO_2

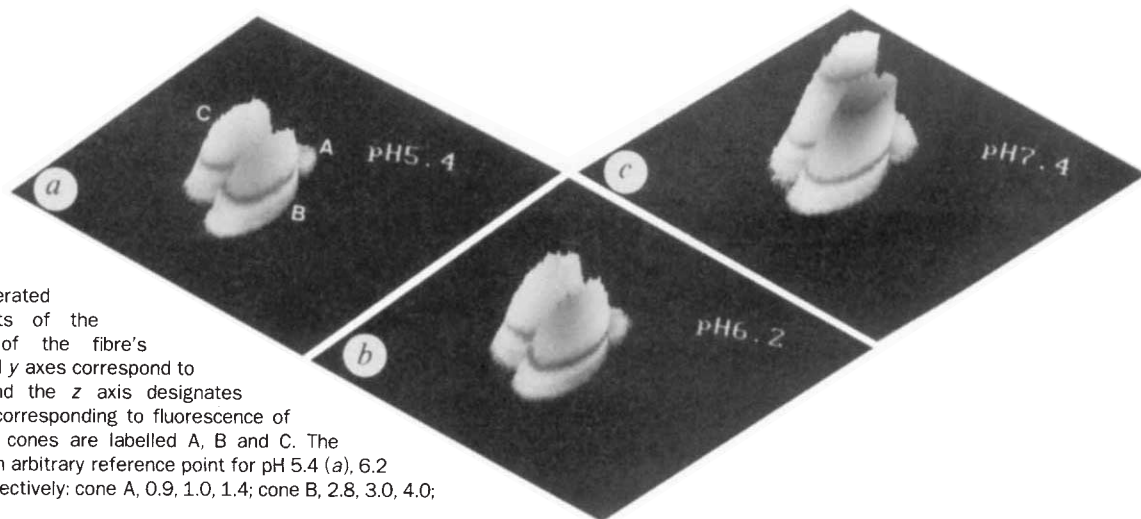


FIG. 3 Computer-generated three-dimensional plots of the fluorescence image of the fibre's proximal end. The x and y axes correspond to spatial coordinates and the z axis designates intensity. The images corresponding to fluorescence of the three pH-sensitive cones are labelled A, B and C. The intensities relative to an arbitrary reference point for pH 5.4 (a), 6.2 (b) and 7.4 (c) are, respectively: cone A, 0.9, 1.0, 1.4; cone B, 2.8, 3.0, 4.0; cone C, 2.9, 3.2, 4.5.

sensor by coating it with a photopolymerized silicone layer. This approach preserves the miniature dimensions of the sensor, and simplifies instrumental design by requiring only a single excitation source, fibre connector and detector. Multi-parameter sensors of this kind should be particularly valuable for *in vivo* medical analysis, where small sensor size is extremely important. □

Received 17 April; accepted 30 July 1991.

1. Collison, M. E. & Meyerhoff, M. E. *Analyt. Chem.* **62**, 425A3-437A (1990).
2. Gehrich, J. L. et al. *IEEE Trans. on Biomedical Engineering* **BM-33**, 117-132 (1986).
3. Collison, M. E., Aebli, G. V., Petty, J. & Meyerhoff, M. E. *Analyt. Chem.* **61**, 2365-2372 (1989).
4. Seitz, R. *CRC crit. Rev. analyt. Chem.* **19**, 135-173 (1988).
5. Wolfbeis, O. S. in *Fiber Optic Chemical Sensors and Biosensors*, Vols 1 & 2 (CRC Press, Boca Raton, 1991).
6. Barnard, S. M. & Walt, D. R. *Science* **251**, 927-929 (1991).
7. Barnard, S. M. & Walt, D. R. *Envir. Sci. Technol.* **25**, 1301-1304 (1991).
8. Wolfbeis, O. S., Weiss, L. J., Leiner, M. J. P. & Ziegler, W. E. *Analyt. chem.* **60**, 2028-2030 (1988).
9. Tsumanuma, T., Tanaka, K., Chigira, S., Sanada, K. & Inada, K. *Proc. SPIE Optical Fibers in Medicine* **906**, 92-96 (1988).
10. Epperson, P. M., Sweedler, J. V., Bilhorn, R. B., Sims, G. R. & Denton, M. B. *Analyt. Chem.* **60**, 327A-335A (1988).
11. Trauthen, B. *Proc. SPIE Optical Fibers in Medicine V* **1201**, 580-583 (1990).
12. Munkholm, C., Walt, D. R., Milanovich, F. P. & Klainer, S. M. *Analyt. Chem.* **58**, 1427-1430 (1986).
13. Munkholm, C., Walt, D. R. & Milanovich, F. P. *Talanta* **35**, 109-112 (1988).
14. Martin, M. M. & Linqvist, L. *J. Luminescence* **10**, 381-390 (1975).

ACKNOWLEDGEMENTS. We thank R. Mansfield for helpful discussions and supplying the imaging fibre, J. Verry and S. Berman for the loan of the Quantex fluorescence imaging system; and F. Milanovich for his help in designing the illuminating-fibre bundle. We gratefully acknowledge the Environmental Protection Agency through Tufts Center for Environmental Management for financial support.

Air-stable alkali-metal colloids and the blue colour in Wurtz syntheses

Robert E. Benfield, R. Harry Cragg, Richard G. Jones* & Anthony C. Swain

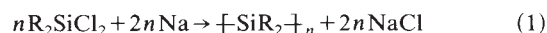
The Centre for Materials Research, The Chemical Laboratory, University of Kent at Canterbury, Canterbury, Kent CT2 7NH, UK

POLYSILANES are used as resists in microlithography and as precursors for silicon carbide ceramics¹⁻³. We have been investigating the formation of polysilanes by the reductive dechlorination polymerization of dichloro-organosilanes with sodium⁴—a modification of the well-known Wurtz reaction, in which organic halides are reductively coupled using sodium to form carbon-carbon single bonds. In all such reactions, after a period of time a blue precipitate is formed, but no characterization of this solid has so far been reported. It has been suggested (without spectroscopic evidence) that the colour in the polysilane synthesis is due to defects in sodium chloride¹, sodium colour centres in sodium chloride^{3,5}, reactive polysilane chain-ends³ or a stabilized organic radical⁶. The first of these suggestions has been tacitly accepted despite the fact that the most common defects in sodium chloride are F centres (anion vacancies with trapped electrons from excess sodium atoms) which give the salt a yellow colour^{7,8}. Here we present spectroscopic results suggesting that the blue colour is due to colloidal alkali-metal particles formed during the reaction. These particles are contained in a matrix composed of an intimate mixture of polymer and alkali-metal halide, and are remarkably stable in air. It is most unusual for colloidal metal particles to be formed under such mild conditions.

It has long been known that the Wurtz reaction is accompanied by the appearance of a blue colour. Indeed it was described by Wurtz in 1855⁹ as a blue crust upon the surface of the sodium metal. There have been a number of studies directed towards establishing the mechanism of the reaction¹⁰, which remains

ill-defined. It is therefore surprising that there are no reports of attempts to characterize the blue colour spectroscopically to identify the species from which it arises and that which contains it.

To obtain a greater understanding of the mechanism of polysilane formation through the Wurtz reaction (scheme 1)



we have studied the blue colour generated in -Si-Si- coupling reactions with sodium and potassium using ultraviolet-visible and electron paramagnetic resonance (EPR) spectroscopy. In a typical reaction, dichloromethylphenylsilane was added to a stoichiometric amount of sodium or potassium dispersion in diethyl ether at room temperature under a dinitrogen or argon atmosphere. Initial attempts to monitor the reaction in a standard 4-mm diameter EPR tube were unsuccessful: in the absence of stirring, the sodium surface soon became coated with a white precipitate, stopping the reaction. The reaction was therefore carried out with stirring in a Schlenk tube equipped with a side-arm EPR tube. After dichlorosilane addition, the supernatant liquid was decanted every ten minutes from the alkali metal dispersion into the EPR tube, the spectrum was obtained, and the liquid returned to the reaction vessel. Although some white precipitation was evident from an early stage, no EPR signal was seen from the liquid and precipitate until 90 minutes after the onset of the reaction. During this time, EPR signals could be observed only from any pieces of bulk sodium in the EPR tube. Most of the precipitate with its characteristic blue colouration then appeared fairly suddenly, and EPR signals were observed from this blue solid. Addition of crown ether to the reaction mixture, while maximizing the yield of high-molecular-weight polymer⁴, had no effect on the spectroscopic measurements. The poly(diphenylsilane)-sodium system had a principal *g*-value of 2.0005 with an additional axial feature at *g* = 2.0112 (Fig. 1a); the linewidth of 8.8 gauss is characteristic of colloidal sodium¹¹⁻¹³. The poly(methylphenylsilane)-potassium system gave a linewidth of ~100 gauss, reflecting the stronger spin-orbit coupling in potassium^{14,15}. These signals were associated only with the blue solid phases. A comparison between the signal obtained from the sodium system, Fig. 1a, and the asymmetric signal for bulk sodium metal¹⁶, Fig. 1b, demonstrates that the signal cannot be attributed to pieces of sodium metal within the precipitate. Our EPR spectra are also inconsistent with those of F-centres, which have quite different linewidths¹⁷, or silane radicals, which produce extensive hyperfine splitting patterns¹⁸.

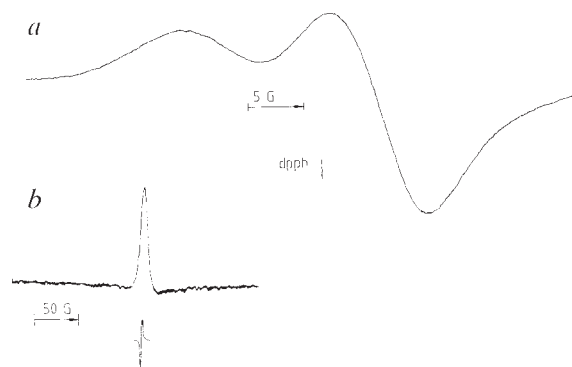


FIG. 1 a, X-band EPR spectrum (JEOL PE-1X) at room temperature of the dried purple solid isolated from the reaction of dichlorodiphenylsilane with sodium in diethylether, recorded in a 10-mm diameter tube. Microwave power 10 mW, modulation 5 gauss, gain 1.8×10^2 ; $g_{\parallel} = 2.0112 \pm 0.0001$; $g_{\perp} = 2.0005 \pm 0.0001$, $\Delta H_{pp} = 8.8 \pm 0.2$ gauss, where pp stands for peak to peak. dpph, diphenylpicrylhydrazyl (standard). b, X-band EPR spectrum of bulk sodium showing the highly asymmetric derivative lineshape. Microwave power 10 mW, modulation 5 gauss, gain 3.2×10^2 ; *g* (corrected for line-shape) = 2.0014 ± 0.0002 , $\Delta H_{pp} = 17 \pm 1$ gauss.

* To whom correspondence should be addressed.

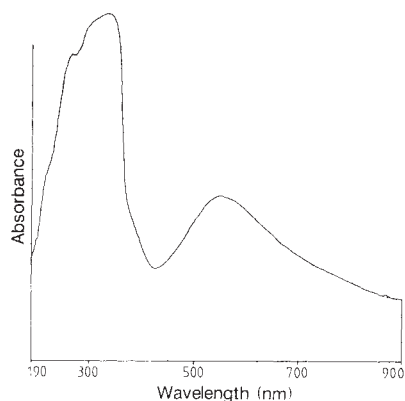


FIG. 2 Ultraviolet-visible diffuse reflectance spectrum (Shimadzu UV-240) of the purple solid from the reaction of dichloromethylphenylsilane with sodium in diethylether. $\lambda_{\max} = 560 \pm 5$ nm; linewidth (FWHM) = 200 ± 5 nm.

The additional feature in the EPR spectrum of Fig. 1a indicates an axially symmetric environment for the sodium colloids. This implies that the colloids are associated with the polymer chain ends rather than just with sodium chloride crystals, which have cubic symmetry. This is consistent with a picture of the mechanism involving polymer chains growing on metal particles consumed during the reaction¹⁹, with residual colloidal metal particles precipitated with an intimate mixture of solid polymer and sodium chloride at the completion of the reaction.

The precipitates were isolated by removal of excess solvent through washing with diethyl ether *in vacuo* followed by evaporation to dryness. Surprisingly, the colour of the precipitate from the sodium systems, which in the dry state appears purple rather than blue, is persistent in air, remaining unchanged for many months. Even in the case of the potassium system, poly(methylphenylsilane)-potassium chloride, the blue colour is stable for some days. EPR spectra of the dried powders were identical to those described above. The colours and EPR signals are rapidly destroyed by normal procedures in the isolation of the polymer²⁰: addition of water, which dissolves the alkali metal halide, or of dry toluene, which dissolves the polymer. This emphasizes the complementary role of the metal halide and polymer in the protective matrix that stabilizes the colloids.

Ultraviolet-visible spectra of the isolated precipitates were obtained as diffuse reflectance spectra. They showed electronic transitions for the polysilanes below 400 nm (ref. 1), and a single broad absorption band centred around 560 nm for the polysilane-sodium chloride powders (see Fig. 2). For the potassium systems there was a similar broad band centred around 750 nm with a shoulder at ~ 540 nm. A comparison of these spectra with published results for alkali metal colloids^{11,21-24} and defect centres^{7,24} shows that the absorptions in the visible region of the spectrum are attributable to colloidal sodium and potassium respectively. They are not consistent with any known atomically dispersed crystalline defect centres such as F, F', H or R₁.

To demonstrate the generality of our observations, we have also studied the products from the reaction of sodium with dibromodiphenylsilane and with 1,10-dibromodecane; the latter reaction forms polyethylene. The ultraviolet-visible reflectance spectra of the dry precipitates from these systems had peak absorptions at $\lambda_{\max} = 600$ nm. This value is consistent with the known absorption characteristics of sodium colloids in sodium bromide²² which has a different dielectric constant from sodium chloride. The $\lambda_{\max}$ values of defects arising from anion vacancies are much more sensitive to changes of halide ion as they depend on the size of the anionic site in the crystal lattice⁷.

All the spectroscopic evidence indicates that the blue-purple precipitates formed in these reactions arise from colloidal alkali metal particles present in the solid phase. From the $\lambda_{\max}$ values and the full widths at half height of the visible absorption bands, particle diameters of ~ 2 nm for sodium and ~ 40 nm for potassium can tentatively be estimated^{21,23}. The shoulder observed in the ultraviolet-visible spectrum of the potassium system might therefore be a quadrupolar absorption in the larger

potassium particles. We are currently investigating the relationship between the colloid sizes and the molecular weight parameters of the product polymers. Transmission electron microscopy has, however, failed to image the colloids because the samples are extremely susceptible to beam damage: the polysilanes undergo main-chain scission (the very property enabling their use as positive-working electron beam resists), and the metal halides are also damaged by electron irradiation at the required voltage²⁵.

Stoichiometric analysis of the poly(methylphenylsilane)-sodium chloride powder has been carried out by a sensitive spectro-titrimetric method²⁶. Samples of the blue powder were hydrolysed with deionized water and the pH of the resulting solutions measured using the absorbance of methyl red indicator monitored at 548 nm. The method was calibrated using ten standard solutions of sodium hydroxide with concentrations spanning the relevant range, and checked with a sample of pure sodium chloride. The colloid concentration was estimated to be 0.4% of the total sodium content; this is too low to be detected by other techniques that were used (²³Na cross-polarized magic-angle spinning NMR and powder X-ray diffraction).

It is significant that these colloidal sodium and potassium particles can be formed directly under such mild conditions. Usual methods for their formation involve high-energy procedures such as metal atom evaporation or ionizing irradiation^{11,25}. The colloids in our polymer systems must be formed directly and not by annealing of F-centre intermediates because the latter reaction requires temperatures above 175 °C (refs 8, 11), much higher than those used in our syntheses. We also emphasize that our results apply to both silicon-silicon and carbon-carbon bond-forming polymerization reactions and suggest that alkali metal colloids may be involved in the mechanism of all Wurtz-type coupling reactions. □

Received 21 May; accepted 30 July 1991.

1. Michl, J. & Miller, R. D. *Chem. Rev.* **89**, 1359-1410 (1989).
2. West, R. J. *organomet. Chem.* **300**, 327-346 (1986).
3. Worsfold, D. G. in *Inorganic and Organometallic Polymers*, ACS Symp. Ser. **360** (eds Allcock, H. R., Wynne, K. J. & Zeldin, M.) 101-111 (American Chemical Society, Washington DC, 1988).
4. Cragg, R. H., Jones, R. G., Swain, A. C. & Webb, S. J. *J. chem. Soc. Chem. Commun.* 1147-1148 (1990).
5. Hengge, E. & Reuter, H. *Naturwissenschaften*, **42**, 514 (1962).
6. Burkhard, C. A. *J. Am. chem. Soc.* **71**, 963-964 (1949).
7. Compton, W. D. & Schulman, J. H. *Color Centers in Solids*, 29: 52-53 (MacMillan, New York, 1962).
8. Doyle, W. T. *Phys. Rev.* **111**, 1067-1072 (1958).
9. Wurtz, A. *Justus Liebig's Ann. Chem.* **96**, 364-375 (1855).
10. March, J. *Advanced Organic Chemistry: Reactions, Mechanism, and Structure*, 2nd edn, 407-412 (McGraw Hill, London 1977).
11. Hughes, A. E. & Jain, S. C. *Adv. Phys.* **28**, 717-828 (1979).
12. Vitol, A. Ya., Kharakhashyan, E. G., Cherkasov, F. G. & Shvarts, K. K. *Fizika Tverd. Tela* **13**, 2133-2135 (1971); (Engl. transl.) *Sov. Phys. Solid St.* **13**, 1787-1789 (1972).
13. Gordon, D. A. *Phys. Rev.* **B13**, 3738-3747 (1976).
14. McMillan, R. C. *J. Phys. Chem. Solids* **25**, 773-775 (1964).
15. Guy, S. C., Edmonds, R. N. & Edwards, P. P. *J. chem. Soc. Faraday II*, **81**, 937-947 (1985).
16. Feher, G. & Kip, A. F. *Phys. Rev.* **98**, 337-348 (1955).
17. Seidel, H. & Wolf, H. C. in *Physics of Color Centres* (ed. Fowler, W. B.) 555 (Academic, New York/London, 1968).
18. Karatsu, T. et al. *J. Am. chem. Soc.* **113**, 2003-2010 (1991).
19. Zeigler, J. N., Harkah, L. A. & Johnson, A. W. *Polymer Preprints* **28**, 424-425 (1987).
20. Trefonas, P. & West, R. *Inorg. Synth.* **25**, 56-60 (1989).
21. Smithard, M. A. & Tran, M. Q. *Helv. Phys. Acta* **46**, 869-888 (1974).
22. Doyle, W. T. *Proc. Phys. Soc.* **75**, 649-663 (1960).
23. Radchenko, I. S. *Optik. Spektrosk.* **25**, 899-902 (1968); (Engl. transl.) *Opt. Spectrosc.* **25**, 500-501 (1968).
24. Scott, A. B., Smith, W. A. & Thompson, M. A. *J. phys. Chem.* **57**, 757-761 (1953).
25. Aso, J., Baba, M. & Ikeda, T. *Jap. J. appl. Phys.* **29**, 308-312 (1990).
26. Kleinschrod, F. G. *Ann. Phys.* **27**, 97-107 (1936).

ACKNOWLEDGEMENTS. We thank S. McIntosh for his experimental assistance. We gratefully acknowledge the support of the Royal Society (R.E.B.) and the SERC (A.C.S.).

Occurrence of small colloids in sea water

Mark L. Wells & Edward D. Goldberg

Scripps Institution of Oceanography, La Jolla, California 92093-0220, USA

COLLOIDAL particles in sea water may be important in the marine chemistry of elements such as carbon and iron^{1,2}. Non-living, submicrometre particles (0.4–1.0 μm) have recently been found to be abundant (10^6 – 10^7 particles ml^{-1}) in the north Pacific³ and off Nova Scotia⁴, but smaller colloidal particles in sea water remain largely uncharacterized. Here we report that marine colloids <120 nm in size are at least three orders of magnitude more abundant than larger submicrometre particles. Moreover, the distribution with depth of these small colloids differs markedly from that reported for their larger counterparts³. This vertical stratification suggests that small colloidal particles in sea water are reactive. Examination by transmission electron microscopy and energy-dispersive X-ray spectroscopy indicates that the colloids are largely organic, although trace metals (Fe, Al, Co) may also be present. The existence of such colloids may contribute to the discrepancy between standard and new high-temperature methods for measuring dissolved organic carbon¹.

Seawater samples were collected in Santa Monica Basin 15 km from shore (33°45.07' N, 118°55.09' W, 908 m depth) during June and September 1990. Samples were obtained with modified Niskin bottles (silicone-rubber coated springs), transferred to acid-cleaned 250-ml Teflon bottles and preserved with HgCl_2 (added to give 1 μM Hg). The samples were wrapped with foil to exclude light and refrigerated for periods up to 1 week before analysis.

Colloids were collected directly on specimen grids for transmission electron microscopy (TEM) by ultracentrifugation, in a modification of the method of Nomizu and Mizuike⁵. Microscope fields ranging in magnification from 42,000 $\times$ to 52,000 $\times$ were selected from different specimen grid openings

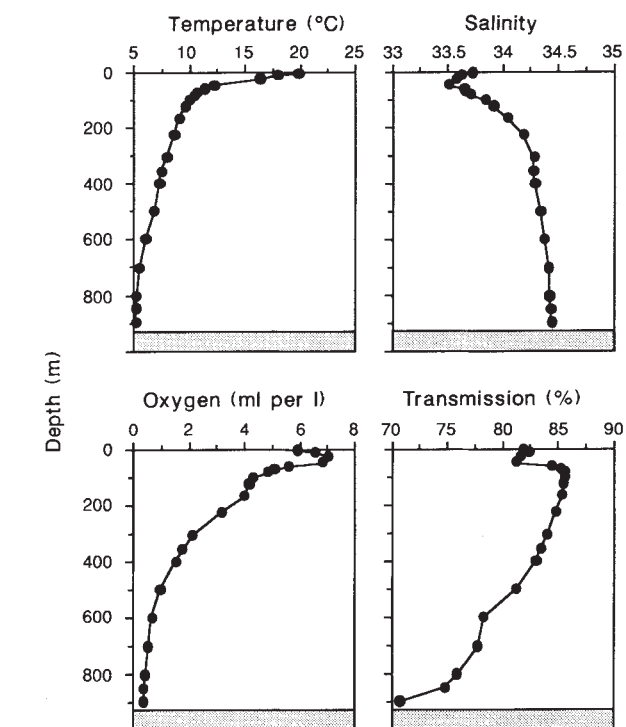


FIG. 2 Physical data for station 305 in Santa Monica basin in July 1990. Comparable data were not available for the September cruise.

and photographed. Colloid numbers and size distributions were then determined by computer image analysis of the digitized TEM micrographs. Elemental compositions of some colloids were determined qualitatively by energy-dispersive X-ray spectroscopy (EDS).

Vertical profiles taken in July and September showed that colloids <120 nm in size were highly stratified. Concentrations remained below detection limits ($\leq 10^4$ particles ml^{-1}) at the surface, increased sharply near the lower thermocline (40–100 m) to $>10^9$ particles ml^{-1} , and decreased again to $<10^4$ particles ml^{-1} below the thermocline (Fig. 1). This distribution is substantially different from that reported for larger (0.4–1.0 μm) colloids in the northwest Pacific, where surface con-

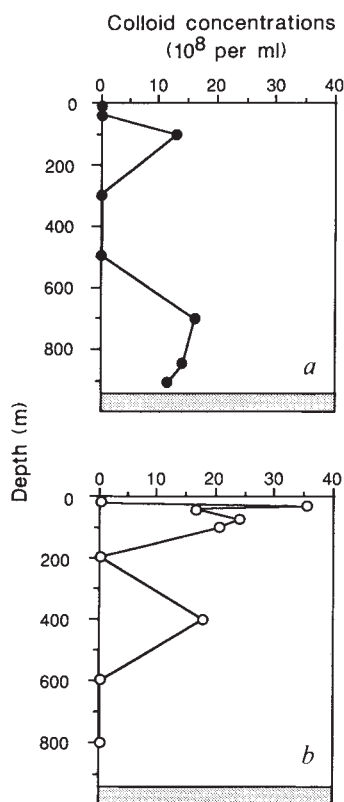


FIG. 1 Vertical profiles of the number of small colloids (5–120 nm) at station 305 in Santa Monica basin (33°45.10' N, 118°55.09' W) on a, 9 July and b, 1 September 1990. Seawater samples (10 ml) were centrifuged (Beckman L7) in a swing bucket rotor (SW41) at 288,000 g for 4.0 h at 25 °C. Settling colloids were collected on 200-mesh copper TEM specimen grids placed in specially constructed holders at the bottom of 13-ml polyallomer centrifuge tubes (Beckman). The grids, coated with carbon-stabilized formvar (Ted Pella®), were first rendered hydrophilic by 60 s of glow discharge under vacuum¹². After centrifugation, the supernatant was removed slowly and the grids rinsed gently with five sequential tube volumes of filtered (0.1 μm Nuclepore) distilled water. The air-dried grids were examined in a Phillips CM 30 transmission electron microscope at 100 kV. Between 200 and 500 colloids were counted in each field, and five to eight fields were photographed from each sample. The detection limit for colloid numbers was constrained by the field area at 42,000 $\times$, the lowest magnification at which 5–10 nm colloids were still visible; two to three colloids per field corresponded to a detection limit of $\sim 1 \times 10^4$ colloids ml^{-1} . Final particle concentrations were calculated taking into account the taper correction for nonparallel particle trajectories during centrifugation¹³. Colloid recovery efficiencies were obtained by adding known amounts of colloidal haematite (~ 600 nm in size) and latex microspheres (88 nm) to sea water; these were $97 \pm 13\%$ and $93 \pm 0.2\%$, respectively. The precision for colloid enumeration was $<10\%$ (± 1 s.d.). Samples and specimen grids were handled under 'clean' conditions in a laminar flow airbench (Class 100); few, if any, colloids were observed in analytical blanks processed in the same way as the samples.

centrations decreased sharply with depth³. The distribution of colloids in deep waters differed markedly between July and September. In July, colloid concentrations in bottom waters (700–910 m) were $>10^9$ particles ml^{-1} , and thus similar to those near the thermocline. By contrast, deep-water concentrations in September were high at 400 m (1.5×10^9 particles ml^{-1}) but remained below the detection limit in bottom waters (Fig. 1*b*). The vertical distribution of small colloids in July did not co-vary with temperature, salinity, dissolved oxygen or light transmission (Fig. 2).

The extreme vertical variation in colloid abundance with depth is surprising, yet we submit that this finding is valid. First, the precision for estimating colloid concentrations on individual specimen grids is better than $\pm 10\%$ (1 s.d. for $n \geq 6$), and triplicate samples taken from the same Niskin bottles yield numbers within this precision. The addition of HgCl_2 does not alter the colloid abundance or size distribution even after five days of storage², longer than the delay between collection and processing for most samples in this study. Successive 4-h centrifugations yielded colloids only after the initial spin. Gentle prefiltering (<155 nm Hg, $0.4 \mu\text{m}$ Nuclepore) of raw sea water before centrifugation did not significantly alter the numbers or size distribution of small colloids recovered (not shown), indicating that they were not artefacts arising from the breakup of larger particles during centrifugation. Vertical distributions of colloids at other stations 30 km and 150 km off the California coast show

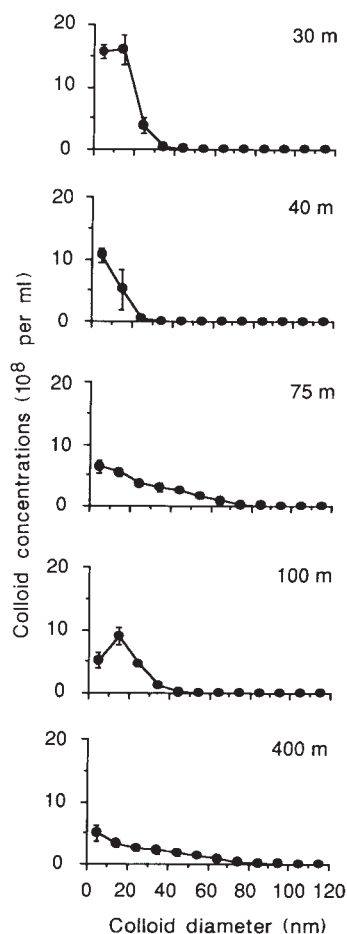


FIG. 3 The size distribution of small colloids collected from five depths at station 305, Santa Monica basin, September 1990. Colloid diameter was taken as twice the minor axis of a theoretical ellipse having the same two-dimensional area as the colloid; size spectra determined this way were not significantly different from those of ruler measurements of the narrowest dimension of colloids in the micrographs. It should be noted that colloid dimensions may have been altered by settling on the formvar film or by dehydration.

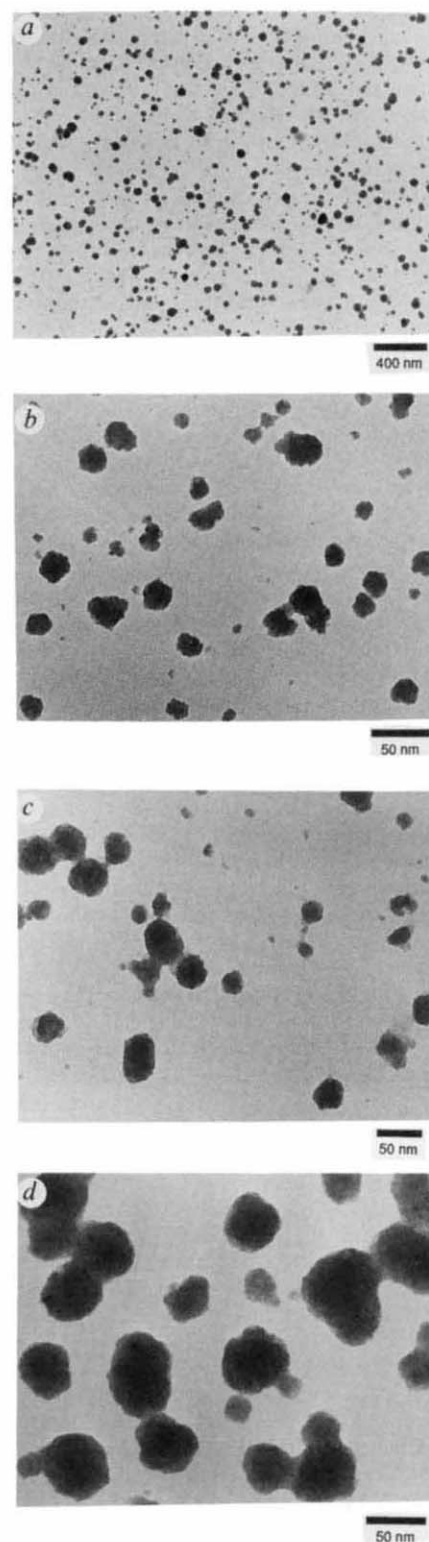


FIG. 4 Transmission electron micrographs from samples collected at station 305, Santa Monica basin in September 1990. The spatial distribution of colloids on the membranes tended to be uniform, as shown for the 75-m sample (*a*). Enlarged views of the colloids from 75 m and 400 m depth are shown in *b* and *c*, respectively. Most of the colloids were globular and rounded, and had low electron opacity, suggesting an organic composition (contrast in these micrographs was enhanced during development). Many appeared to be aggregates of smaller granules 2–10 nm in size, some of which were darker in contrast indicating an inhomogeneous colloid composition. *d*, A similar inhomogeneous, aggregative morphology is seen clearly in colloids from 850 m in July 1990.

common features, with high abundances in the region of the thermocline and numbers below detection in deeper waters (M.L.W. and E.D.G., manuscript in preparation). Colloids were also abundant in surface (1 m) sea water 2 km offshore of the Scripps Institution of Oceanography from February through April² when upwelling conditions predominated. Colloid numbers fell below the detection limit, however, in July and August when upwelling ceased and the water column became stratified (as were conditions in Santa Monica basin during the time of sampling). It should be noted that values reported here as $<10^4$ particles ml^{-1} do not indicate that colloids of $<0.1 \mu\text{m}$ are completely absent, but that their small numbers prevented quantification.

The size spectrum below 120 nm diameter showed an increase in colloid numbers with decreasing size that was nearly logarithmic in some cases (Fig. 3). This size relationship may be considerably more pronounced than indicated here because the smallest size classes ($<40 \text{ nm}$) may not have been collected in representative numbers by 4 h of centrifugation. Logarithmic size distributions are very common for marine particles $>1.0 \mu\text{m}$ (ref. 6) and these findings indicate that this general relationship extends into the smallest of the colloidal phases.

The nature and composition of these small colloids can be inferred from high-resolution TEM images and EDS analyses. The $<120 \text{ nm}$ size class included marine viruses and, less frequently, inorganic crystalline materials. But by far the most abundant components of this size class are discrete, relatively diffuse colloids having low electron opacity (Fig. 4a-d). The colloids tended to be rounded in shape and many appeared to be aggregates of smaller granules $\sim 2\text{--}10 \text{ nm}$ in size. Microfibrils, such as those found abundantly in fresh-water lakes^{7,8}, were not found. The low electron opacity of these colloids suggests that they are largely organic, although dark (high-contrast) granules were often observed within them, indicating an inhomogeneous composition (Fig. 4b-d). The colloid morphology was similar in shallow and deep waters, perhaps indicating a common source or mechanism of their formation.

In appearance, these colloids bear a striking resemblance to transmission electron micrographs of soil-derived fulvic acids, which also show a granular, aggregated morphology^{9,10}. A granule size of $2\text{--}5 \text{ nm}$ corresponds to $\sim 10,000$ daltons^{9,11}, which is well within the range for dissolved organic carbon (DOC) fractions measured in sea water¹. EDS analysis of colloids from surface and deep waters yielded rather patternless spectra, which is consistent with a largely organic matrix. Most of the colloids examined fell below the instrumental size limit suggested for 'quantitative' EDS analyses ($\sim 100 \text{ nm}$). All of the particles examined, however, were slightly enriched in Fe, Co, Si, Al and O (19 particles examined). In some cases, Cr, Ca, Ni and V were also detected.

The vertical stratification (and therefore nonconservative behaviour) of these small colloids, coupled with their apparent organic nature, suggest they are derived primarily from biological processes and that they have short residence times in sea water (although the colloid suspensions seem to be stable² for at least five days in refrigerated sea water preserved with HgCl_2). It is not clear what processes produce the depth distributions found. The occurrence of these colloids in waters off the Scripps Institution of Oceanography² and on the outer continental slope off Point Conception, California (M.L.W. and E.D.G., manuscript in preparation) indicates that they are dispersed widely in marine waters. If their composition is typical of biological debris, they could account for up to $\sim 10\%$ of DOC in these waters (assuming spherical shape, a mean density of 1.1 g cm^{-3} , 50% C by weight and 1.0 mg l^{-1} DOC), and may contribute to the discrepancy between standard and new high-temperature methods for DOC determinations¹. Furthermore, the apparent close association of metals with these colloids suggests that they may play an important part in the transport and fate of trace elements in sea water. □

Received 7 February; accepted 29 July 1991.

1. Sugimura, Y. & Suzuki, Y. *Mar. Chem.* **24**, 105–131 (1988).
2. Wells, M. L. & Goldberg, E. D. *Mar. Chem.* (in the press).
3. Koike, I., Hara, S., Terauchi, K. & Kogure, K. *Nature* **345**, 242–244 (1990).
4. Longhurst, A. *et al. Deep-Sea Res.* (submitted).
5. Nomizu, T. & Mizuike, A. *Mikrochim. Acta* **1**, 65–72 (1986).
6. McCave, I. N. *Deep-Sea Res.* **31**, 329–352 (1984).
7. Leppard, G. G., Massalski, A. & Lean, D. R. S. *Protoplasma* **92**, 289–309 (1977).
8. Leppard, G. G. *Arch. Hydrobiol.* **101**, 521–530 (1984); *Water Res.* **20**, 697–702 (1986).
9. Leppard, G. G., Buffle, J. & Baudat, R. *Water Res.* **20**, 185–196 (1986).
10. Schnitzer, M. & Kodama, H. *Geoderma* **13**, 279–287 (1975).
11. Williams, P. M. in *Plankton Dynamics of the Southern California Bight* (ed. Eppley, D.) 53–83 (Springer, New York, 1986).
12. Hayat, M. A. *Basic Techniques for Transmission Electron Microscopy* 234–235 (Academic, Orlando, 1986).
13. Mathews, J. & Buthala, D. A. *J. Virology* **5**, 598–603 (1970).

ACKNOWLEDGEMENTS. We thank G. Hemingway, E. Ringer, and W. Cochran for collecting water samples on our behalf. We also thank P. Buseck and W. Su for assistance with EDS measurements. This study was funded by ONR.

Structural transformation of quartz at high pressures

N. Binggeli & James R. Chelikowsky

Department of Chemical Engineering and Materials Science and Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, Minnesota 55455, USA

THE behaviour of α -quartz, one of the most common tetrahedrally coordinated silica polymorphs, at high pressures has been the subject of several experimental studies. At room temperature, gradual pressure-induced amorphization is observed (at about $25\text{--}35 \text{ GPa}$)^{1,2}, followed at higher pressures (above 60 GPa) by a transformation to a crystalline octahedrally coordinated 'rutile-like' structure³. The driving force for these transformations is not well understood. We report here first-principles calculations of the electronic and structural properties of the high-pressure structure of α -quartz, which show a pressure-induced transformation of the oxygen sublattice to a body-centred cubic structure. The main features of cubic packing are present at 30 GPa , and the ideal form is achieved at about 60 GPa . The formation of the cubic sublattice facilitates structural transformations involving a change in Si coordination. In the oxygen body-centred cubic lattice, a small displacement of Si ions along the empty channels of the structure is sufficient to transform α -quartz either to a structure with mixed fourfold/sixfold coordination or, for a larger silicon displacement, to a purely sixfold-coordinated structure. This description is consistent with the evidence of intermediate crystalline phases with mixed Si coordination above the amorphization pressure in molecular-dynamics simulations⁴, and can also explain the transformation to a sixfold crystalline structure at higher pressure.

Silica, an important constituent of the Earth's crust, exists in a number of allotropic forms. Among the various silica polymorphs, α -quartz is the stable structure at room temperature and at pressure below 3 GPa . At higher pressures, this structure persists as a metastable phase which gradually transforms to an amorphous state and then to a crystalline structure in which Si is sixfold coordinated. The nature of the amorphous state is not known. Molecular-dynamics simulations based on a two-body potential by Tsuneyuki *et al.*⁴ suggest that this phase may contain sixfold and fourfold coordinated Si atoms. In some of these simulations, the authors found that quartz transforms to crystalline phases with mixed fourfold/sixfold coordinations above 30 GPa . In other simulations, they found a transformation to a sixfold coordinated phase above 30 GPa , and they propose that in high-pressure experiments compression of a macroscopic sample would produce coexistence of domains of the different phases.

The model of polyhedral tilting⁵ generally used to describe the structural changes in silica polymorphs consisting of corner-

connected $\text{Si}(\text{O}_4)_{1/2}$ tetrahedra cannot explain⁶ the large deformation observed in the range 0–11 GPa of the oxygen sublattice in α -quartz. Sowa has suggested⁶ that a change in the oxygen packing towards a body-centred cubic (bcc) stacking was responsible for the observed deformation. This transformation was extrapolated from the trends observed during compression below 11 GPa of α -quartz and its low-pressure isomorphic counterpart α - GeO_2 . Here we examine this model for the high-pressure structure of α -quartz by means of first-principles calculations. Unlike the molecular-dynamics simulation⁴, this study is based on the true quantum-mechanical interactions and electronic structure. The calculations are performed within the local-density approximation to the density functional theory using the pseudopotential plane-wave approach. The details of the calculations have been described elsewhere^{7,8}. In the geometry optimization we retain the α -quartz symmetry. We have checked that this structure corresponds to a stable minimum also at high pressures.

The α -quartz structure has hexagonal D_3^4 symmetry, and is defined by two lattice constants (c , a) and four internal parameters (u , x , y , z). The parameter u and the set of parameters (x , y , z) define the positions in the unit cell of the three Si and six O ions, respectively⁸. The structural parameters derived from the model of the ideal bcc packing of oxygen are $x = y = 1/3$, $z = 1/12$ and $c/a = \sqrt{3/2} \approx 1.225$. Assuming ideally centred $\text{Si}(\text{O}_4)_{1/2}$ tetrahedra, in addition we have $u = 5/12 \approx 0.417$. The structural parameters obtained from the first-principles calculations are found to change from $x = 0.418$, $y = 0.274$, $z = 0.118$, $u = 0.469$ and $c/a = 1.124$ at zero pressure to $x = 0.350$, $y = 0.339$,

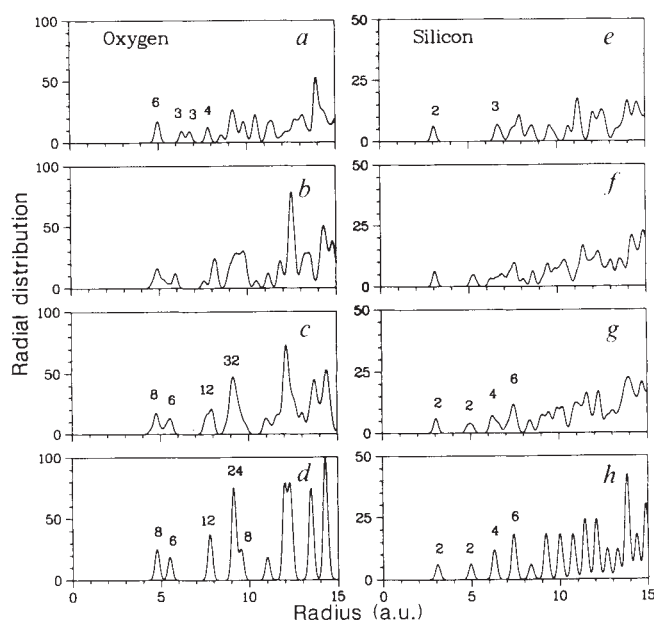


FIG. 2 Radial distribution functions for the oxygen ions (left-hand panels) and silicon ions (right-hand panels) in α -quartz. *a*, *e*, At zero pressure; *b*, *f*, at 30 GPa; *c*, *g*, at 60 GPa; *d*, *h*, radial distribution functions in the model oxygen bcc structure. In *h*, the Si ions are on ideal tetrahedral sites in the bcc lattice. The origin is at an oxygen site, and the densities have been convoluted with a gaussian broadening function with a full width at half maximum of 0.3 a.u. For the ideal oxygen bcc structure (*d* and *h*), we used the same unit-cell volume as in *c* and *g*, where the compression is $V/V_0 = 0.652$.

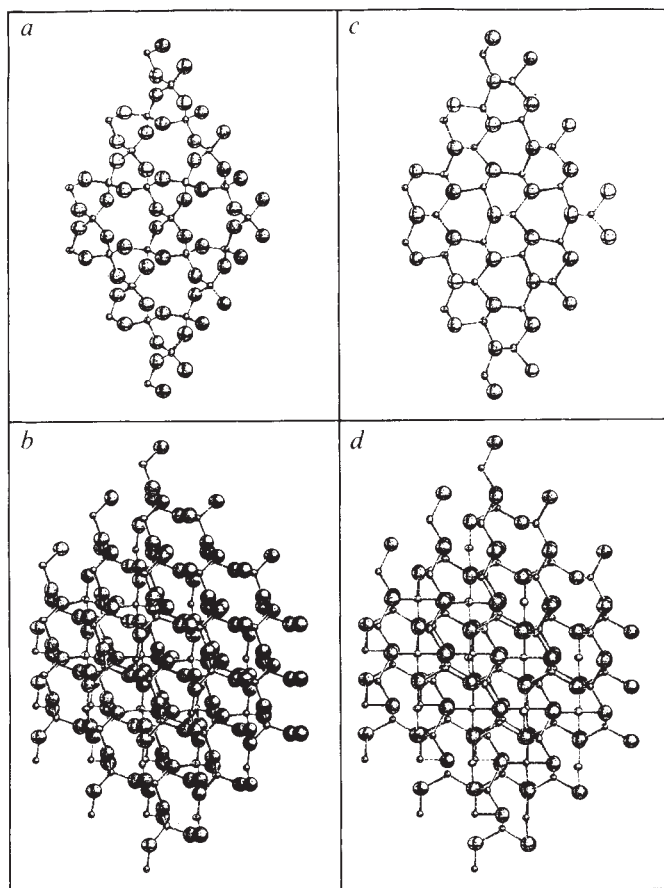


FIG. 1 Fragment of α -quartz at zero pressure (left-hand panels) and at 60 GPa (right-hand panels). *a*, *c*, As seen from the c -axis, corresponding to the $[111]$ direction of the bcc model cell. *b*, *d*, As seen from the $[100]$ direction of the bcc cell. The small spheres represent silicon ions and the large spheres oxygen ions.

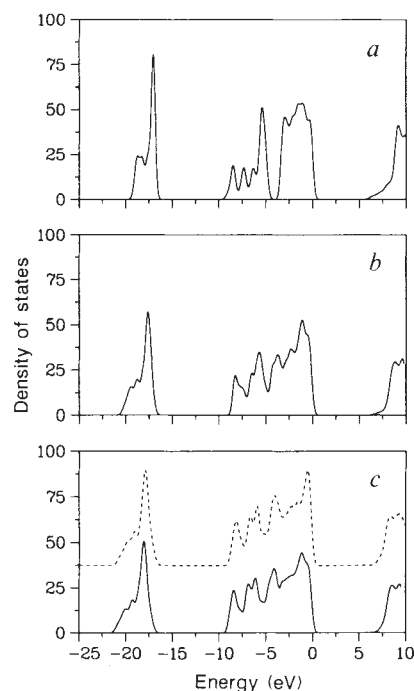


FIG. 3 Calculated density of states for α -quartz at *a*, zero pressure, *b*, 30 and *c*, 60 GPa. The zero of the energy scale corresponds to the valence-band maximum. In *c*, the dashed line shows the density of states calculated for the ideal bcc packing of the oxygen ions (the volume is the same as calculated at 60 GPa, that is, $V/V_0 = 0.652$, and we use the optimized silicon parameter $u = 0.410$). The calculated densities have been convoluted with a gaussian broadening function of full width at half maximum 0.5 eV.

$z = 0.077$, $u = 0.406$ and $c/a = 1.149$ for a compression $V/V_0 = 0.652$, corresponding to a pressure of ~ 60 GPa (ref. 8). The change in the internal parameters at high pressures agrees well with the prediction of the bcc model. The c/a ratio, however, is smaller than expected for the ideal bcc packing of oxygen. To identify bcc packing we can explore the structure along the high-symmetry directions of the ideal cubic framework. In Fig. 1, a section of the structure obtained at zero pressure (panels *a* and *b*) and 60 GPa (panels *c* and *d*) is projected along the c -axis of the α -quartz hexagonal unit cell (panels *a* and *c*), which corresponds to the $[111]$ direction of the model bcc cell, and along the $[100]$ direction of the bcc cell (panels *b* and *d*). The oxygen structure, in Fig. 1*c*, exactly corresponds to a bcc lattice seen along the diagonal of the cube. In Fig. 1*d*, the oxygen sublattice also reproduces the characteristic (100) face of a bcc structure, except for a small misalignment of the O ions along $[100]$.

In the four panels on the left-hand side of Fig. 2, we compare the radial distribution functions of the oxygen ions in α -quartz at ambient pressure, 30 GPa and 60 GPa, and in the model bcc structure. On the right-hand side, we show the corresponding radial distribution functions for silicon. The unit-cell volume for the model case (Fig. 2*d* and *h*) was chosen to be identical to that at 60 GPa (*c* and *g*), and the Si ions were placed on ideal tetrahedral sites in the bcc lattice. There is a striking similarity between the distribution function for O in the α -quartz structure at ~ 60 GPa (*c*) and in the model structure (*d*). The two structures have virtually the same radial distribution function for the nearest-neighbour oxygen shells shown in the figure. The number of oxygen ions in the first three shells, which is (6, 3, 3) at low pressure (*a*), becomes (8, 6, 12) at high pressure (*c*), which corresponds exactly to the distribution of the model bcc structure (*d*). The analogy between the two structures also appears in the silicon distribution functions in Fig. 2*g* and *h*. We note that the peak at 5 a.u., which is absent in α -quartz at low pressure (Fig. 2*e*), appears instead at the right position at ~ 30 GPa (Fig. 2*f*), when the oxygen sublattice approaches the bcc structure. Although the c/a ratio at 60 GPa is smaller than expected for the ideal bcc packing, the overall changes with pressure in the radial distribution functions clearly indicate that the structure becomes increasingly similar to the bcc model.

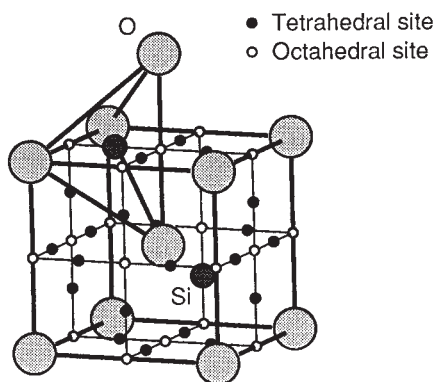


FIG. 4 Octahedral (○) and tetrahedral (●) sites in the oxygen bcc lattice. The large shaded circles correspond to the O ions, and the small shaded circles indicate tetrahedral sites occupied by Si ions in the high-pressure structure of α -quartz. The Si paths to the nearby octahedral sites and also to neighbouring empty tetrahedral sites run along tunnels between the columns of oxygen ions.

The trends in the pressure-induced changes of the α -quartz electronic structure constitute an important probe of the bcc model. In Fig. 3, we show the density of states of α -quartz calculated at zero pressure, 30 and 60 GPa, and for the model bcc structure. In Fig. 3*c*, for the model structure, we have used the same unit-cell volume as obtained from the calculations at 60 GPa, and the optimized parameter $u = 0.410$ for Si. This value is very close to the value expected when Si ions occupy ideal tetrahedral sites in the bcc lattice ($u \approx 0.417$). At low pressure, the α -quartz density of states (Fig. 3*a*) is very different from that found for bcc packing of oxygen. At 60 GPa, instead, features in the density of states for α -quartz are essentially identical to those obtained for the model bcc packing. In particular, we note fourfold structure in the density of states of the upper valence Si-O($2p$) bonding and O($2p$) nonbonding band (from 0 to -10 eV), which is a characteristic of the model bcc structure (also at zero pressure). This indicates that the oxygen bcc model can be considered as the high-pressure limit on the basis of the evolution of the electronic structure.

The ideal bcc packing is achieved at pressures significantly larger than the amorphization pressure. The trends with pressure (Figs 2 and 3) show, however, that the main features associated with the bcc packing already occur at pressures of ~ 30 GPa. Different mechanisms associated with the deformation towards the ideal bcc limit may play a key part in the transition to the amorphous state. A bending of the Si-O-Si angles below 120° is necessary to achieve the ideal oxygen bcc structure². According to molecular calculations⁹, this deformation is associated with a large increase of strain energy. To relieve such strain the crystal could transform to an amorphous phase. Another explanation, which would be more consistent with the recently reported transformation³ of α -quartz to a sixfold coordinated rutile-like structure at ~ 70 GPa, is that the oxygen bcc packing, even in a nonideal form, may promote the diffusion of the Si ions. The bcc cell has 12 equivalent tetrahedral sites and 6 octahedral sites, as shown in Fig. 4. Only one of the tetrahedral sites (two half-sites at the surface of the cell in Fig. 4) is occupied in the α -quartz structure. Diffusion of Si to the nearby octahedral sites or between the tetrahedral sites along tunnels formed by the columns of oxygen ions (see Fig. 4) may be a key process in the transformation to the amorphous state. In fact, migration of two of the three equivalent Si, in the α -quartz unit cell, from the tetrahedral to one of the neighbouring octahedral sites in the bcc lattice leads to the three crystalline phases of mixed coordination found in the simulations performed by Tsuneyuki *et al.*⁴. The oxygen bcc lattice is only slightly affected by the displacement of the Si (compression of a few per cent along one of the three equivalent directions in the α -quartz structure). We note that one of the independent O-Si-O angles, which is equal to $\sim 110^\circ$ in α -quartz at low pressure, increases to $\sim 127^\circ$ in the bcc case, providing a path for enhanced Si diffusion to the octahedral sites or between the tetrahedral sites. □

Received 3 June; accepted 30 July 1991.

1. Hemley, R. J., Jephcoat, A. P., Mao, H. K., Ming, L. C. & Manghnani, M. H. *Nature* **334**, 52-54 (1988).
2. Hazen, R. M., Finger, L. W., Hemley, R. J. & Mao, H. K. *Solid State Commun.* **72**, 507-511 (1989).
3. Tsuchida, Y. & Yagi, T. *Nature* **347**, 267-269 (1990).
4. Tsuneyuki, S., Matsui, Y., Aoki, H. & Tsukada, M. *Nature* **339**, 209-211 (1989).
5. Hazen, R. M. & Finger, L. W. *Comparative Crystal Chemistry* (Wiley, New York, 1982).
6. Sowa, H. Z. *Kristallogr.* **184**, 257-268 (1988).
7. Cheikowsky, J. R., King, H. J. Jr, Troullier, N., Martins, J. L. & Glinemann, J. *Phys. Rev. Lett.* **65**, 3309-3312 (1990).
8. Cheikowsky, J. R., Troullier, N., Martins, J. L. & King, H. Jr *Phys. Rev. B* (in the press).
9. Geisinger, K. L., Gibbs, G. V. & Navrotsky, A. *Phys. Chem. Minerals* **11**, 266-283 (1985).

ACKNOWLEDGEMENTS. We thank H. E. King Jr and J. L. Martins for discussions. This work was supported by the US Department of Energy of the Office of Basic Energy Sciences (Division of Materials Research), and by the Minnesota Supercomputer Institute. N.B. acknowledges support from the Swiss National Science Foundation.

Experimental detection of rapid evolutionary response in natural lizard populations

A. Malhotra & R. S. Thorpe

Department of Zoology, University of Aberdeen, Tillydrone Avenue, Scotland AB9 2TN, UK

MANY studies of geographical variation within species¹⁻⁷, including those using numerical hypothesis tests³⁻¹⁰, have demonstrated a relationship between patterns of phenotypic and environmental variation. But relatively few rigorously tested direct demonstrations of current selection in natural populations exist¹¹⁻¹⁴. Here we present evidence of a rapid response to selection from a field manipulation of the Dominican lizard, *Anolis oculatus*. There is considerable altitudinal and longitudinal variation in climate and vegetation on the island of Dominica¹⁵. We have recorded complex patterns of geographic variation in morphology (body size and shape, colour pattern and scalation), which we have shown to correlate (both univariately and multivariately) with these patterns of ecological variation¹⁶ by numerical hypothesis testing^{10,17,18}. Populations of several ecotypes of the species were translocated into large-scale experimental enclosures, and monitored over a period of two months. The magnitude of the difference in multivariate morphology between survivors and non-survivors within each enclosure was found to correlate with the magnitude of the difference between the ecological conditions of the enclosure site and the original habitats. Similar relationships were found for three indices of fitness of survivors.

The manipulative field experiment described here was designed to test the adaptive nature of morphological variation. We constructed four large-scale lizard-proof enclosures and translocated samples of four 'source' populations, representing the four main ecotypes (Fig. 1) into these enclosures. One enclosure contained a 'resident' control, which was subjected to the same procedures as translocated 'foreign' ecotypes. Before numbering and releasing into the appropriate enclosure, 10 morphological characters (including body proportions, scalation and colour pattern characters) were recorded from each lizard. This multivariate phenotypic profile was later used to compare morphology of survivors and non-survivors (surviving lizards were not remeasured for this profile). Body weight was also recorded (to the nearest 0.01 g).

The enclosures were constructed in May 1990, stocked in June and July (at the start of the wet season), and populations were monitored in September 1990 after a period of about two months had elapsed. Monitoring consisted of one to two weeks in which all enclosures were intensively searched concurrently. Body weight and snout-vent length of each lizard was measured on capture. They were then marked near the base of the tail with

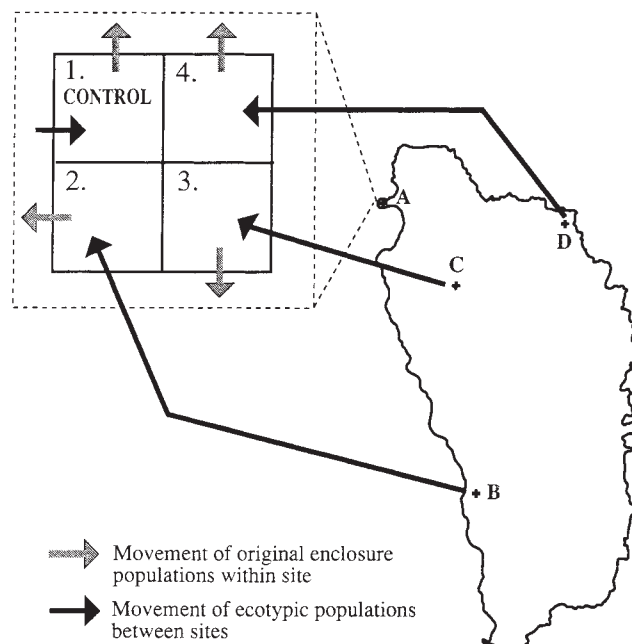


FIG. 1 Schematic representation of the experimental design. The lizard-proof enclosures were based loosely on the design described in ref. 20. The areas were enclosed by a 1-m-high fence (agricultural polyethylene netting of small mesh size), dug into the ground at the bottom and topped by a 30-cm-wide overhang of smooth plastic, with a 2-m gap to the nearest vegetation (from ground level to the canopy) on either side. Construction details will be published elsewhere. The actual dimensions of each enclosure were measured, and their area calculated. Anole biomass per unit area of habitat was determined by capturing and weighing all original inhabitants, and averaged over all enclosures. They were then stocked at the same level with animals from populations representative of the different ecotypes. Control populations were also moved from adjacent habitat into the control enclosure. The morphological characters defining the four ecotypes and their representative populations, were decided on the basis of a previous multivariate investigation of geographic variation in *Anolis oculatus* (manuscript in preparation). The North Caribbean ecotype, represented in our study by a population from the Cabrits (A; also the site of the experimental enclosures), occupies low scrub-like woodland in the most arid part of the island. The South Caribbean ecotype from the less arid southern west coast was represented by a sample from Morne Daniel (B). The Montane ecotype was represented by Syndicate Estate (C), a rainforest site at an altitude of 660 m. Lastly, the Atlantic ecotype was taken from dense, wind-blown littoral woodland at Eden Estate (D). Animals were weighed, sexed and numbered. To minimize stress, animals were anaesthetized before being processed. They were released into the enclosures as soon after capture as possible.

TABLE 1 Means of three indices of fitness of surviving individuals in each enclosure

Ecotypes in enclosures*	Enclosure area (m ²)	Biomass of site (g m ⁻²)	N (start)	N (end)	Growth rate (mm d ⁻¹)	Change in condition	Rate of weight gain (g d ⁻¹)	D ²		P	
								M	F	M	F
1. N. Caribbean	131.85	2.94	103	58	0.080	24.74	0.027	1.83	0.49	NS	NS
2. S. Caribbean	100.35	2.94	63	28	0.045	-2.56	0.010	1.52	0.76	NS	NS
3. Montane	109.35	2.94	38	19	0.010	-48.20	-0.033	10.95	8.37	<0.01	<0.001
4. Atlantic	143.10	2.94	80	52	0.042	11.99	0.014	1.68	1.96	NS	NS

The difference between the means in all cases is highly significant (ANOVA, $P < 0.01$). Condition is defined as SVL^3/WT , Mahalanobis distances (D^2) between survivors and non-survivors, and associated probabilities (P), are given for males (M) and females (F) separately. N , number of specimens in enclosure; NS, not significant.

* Numbers are enclosure numbers as in Fig. 1.

a permanent marker pen, and released. Searching was only abandoned when two consecutive searches failed to reveal any additional unmarked lizards.

The hypothesis that ecotypic variation is caused by natural selection for differing environmental conditions was tested in two ways. If selection is strong enough, we should detect a difference between the morphology of animals that survived and those that died within any one generation of the respective ecotypes. Furthermore, the intensity of this selection should be related to the magnitude of relative ecological differences experienced by the translocated ecotypes. Also, the surviving animals of translocated ecotypes should show detectably lower 'fitness' (the implicit assumption being that the indices of fitness used here translate into relative lifetime fitness as defined by Endler¹²) compared with the control 'resident' ecotype.

We tested for a difference in the morphology of survivors and nonsurvivors using two methods. Firstly, a 3-way multivariate analysis of variance (MANOVA) was used to examine overall differences in survival between ecotypes (Montane, North Caribbean, South Caribbean, Atlantic, as defined previously¹⁶). The model included interactions between sex, survival and ecotype. The interaction between survival and ecotype reveals whether the magnitude or morphological difference between survivors and non-survivors varies according to ecotype (that is, varying selection intensity), and the results show a highly significant difference does exist ($P < 0.001$). A canonical variate analysis was performed on all groups (4 ecotypes $\times$ 2 sexes $\times$ survival factor = 16 groups) and the multivariate distance (Mahalanobis, D^2) between the morphology of survivors and nonsurvivors of each ecotype was obtained.

A plot of the magnitude of the morphological difference between survivors and non-survivors (D^2), against the magnitude of the difference between ecological conditions at the enclosure site and the original habitats shows a significant correlation (Fig. 2). This indicates that there has been a selective deletion of certain phenotypes over a very short time (2 months), and that the intensity of this selection is significantly related to the change in ecological conditions that the ecotypes have experienced. More extreme ecological changes (as experienced by the Montane ecotype) results in highly significant differences

(Table 1) between morphology of survivors and nonsurvivors, in spite of the low sample sizes.

Three indices of fitness of surviving individuals were also calculated; these were growth rate (mm per day), weight gain (g per day) and change in condition (where condition is defined as SVL^3/WT , SVL being the snout-vent length, and WT the body weight). These were calculated for males only, to avoid confusion arising from differences in reproductive effort in females. Table 1 shows that the fitness of the control ecotype is better than that of the translocated ecotypes for all three indices. Among translocated ecotypes, the coastal ecotypes show higher fitness than the montane ecotype.

There is much discussion of the role and mode of action of natural selection in evolution¹⁹. This experiment, designed to run over a long-term period, has unexpectedly demonstrated that significant mortality selection can occur over a very short time-scale in perturbed populations. By showing that the intensity of selection acting on the translocated ecotypes is correlated with the magnitude of ecological change, we provide direct evidence that the high degree of morphological variation in this species is maintained by selection for different habitats, in the absence of barriers to gene flow between adjacent populations. □

Received 3 July; accepted 9 August 1991.

1. Johnston, R. F. & Selander, R. K. *Evolution* **26**, 20–31 (1973).
2. Kahler, A. L., Allard, R. W., Krzakowa, M., Wehrhahn, C. F. & Nevo, E. *Theor. appl. Genet.* **56**, 31–47 (1980).
3. Thorpe, R. S. & Baez, M. *Evolution* **41**, 256–268 (1987).
4. Brown, R. P., Thorpe, R. S. & Baez, M. *Nature* **352**, 60–62 (1991).
5. Brown, R. P. & Thorpe, R. S. *Biol. J. Linn. Soc.* (in the press).
6. Thorpe, R. S. & Brown, R. P. *Biol. J. Linn. Soc.* **38**, 303–322 (1989).
7. Thorpe, R. S. & Brown, R. P. *Herpetologica* **47**, 28–37 (1991).
8. Thorpe, R. S. *Syst. Zool.* (in the press).
9. Dillon, R. T. *Syst. Zool.* **33**, 69–82 (1984).
10. Douglas, M. E. & Endler, J. A. *J. theor. Biol.* **99**, 777–795 (1982).
11. Endler, J. A. *Evolution* **34**, 76–91 (1980).
12. Endler, J. A. *Natural Selection in the Wild* (Princeton University Press, New Jersey, 1986).
13. Halkka, O. & Raatikainen, M. *Hereditas* **80**, 27–34 (1975).
14. Knights, R. W. *Genetica* **50**, 51–60 (1979).
15. Hodge, W. H. *Georg. Rev.* **33**, 349–375 (1943).
16. Malhotra, A. & Thorpe, R. S. *J. evol. Biol.* **4**, 321–335 (1991).
17. Sokal, R. R. *Syst. Zool.* **28**, 227–232 (1979).
18. Dietz, E. J. *Syst. Zool.* **32**, 21–26 (1983).
19. Endler, J. A. in *Speciation and its Consequences* (eds Otte, D. & Endler, J. A.) 625–648 (Sinauer, Massachusetts, 1989).
20. Pacala, S. W., Rummel, J. & Roughgarden, J. *J. Herpetol.* **17**, 94–97 (1983).

ACKNOWLEDGEMENTS. The authors contributed equally to this research. We thank F. Gregoire (Director, Forestry and Wildlife Department, Commonwealth of Dominica) for permission to site the enclosures at the Cabrits, the Geest Line for shipping materials, and P. Pierre (Warden, Fort Shirley, Cabrits National Park). This research was supported by the SERC (A.M.), the University of Aberdeen Research Committee, The Carnegie Trust and The Bonhote Trust (R.S.T.).

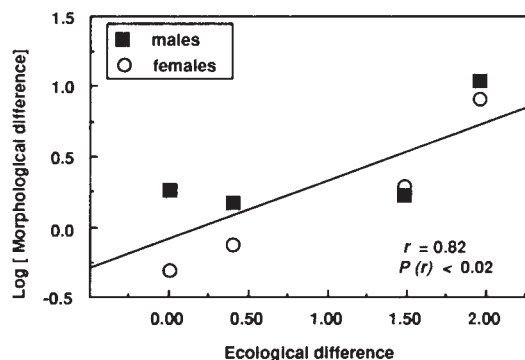


FIG. 2 The relationship between the change in ecological conditions experienced by the ecotypes, and morphological distance between survivors and non-survivors. Ecological difference is derived from a taxonomic distance matrix based on six ecological variables (namely average annual rainfall, average annual temperature, altitude and occurrence of three vegetation types). The habitats of the translocated ecotypes increasingly differ from that of the enclosure site (North Caribbean) in the following order: South Caribbean, Atlantic, Montane. The difference in morphology of survivors and non-survivors is represented by the Mahalanobis D^2 from the canonical variate analysis (see text). The morphological characters used were snout width; length of the fourth toe of the hind foot; lower leg length; the number of scale rows at mid-body; the number of scales between the supra-orbital semicircles; the degree of enlargement of lateral white scales; the number of white lateral spots covering more than 20 scales; dewlap hue and ventral hue.

A cell autonomous function of *Brachyury* in *T/T* embryonic stem cell chimaeras

P. Rashbass, L. A. Cooke, B. G. Herrmann* & R. S. P. Beddington

AFRC Centre for Genome Research, King's Buildings, Edinburgh EH9 3JQ, UK

* Max-Planck-Institute für Entwicklungsbiologie, Spemannstrasse 35, D-7400 Tübingen, Germany

DEVELOPMENTAL genetics has shown that the *Brachyury* (*T*) gene has a key role in mesoderm formation during gastrulation in the mouse^{1,2}. Homozygous embryos (Fig. 1) have a defective allantois, degenerate or absent notochord and disrupted primitive streak and node. The neural tube is kinked and somite formation interrupted^{3–6}. The *T* gene has been cloned⁷ and is expressed during the early stages of gastrulation, being restricted to the primitive streak region, nascent mesoderm and notochord⁸. Neither

TABLE 1 Development and frequency of chimaerism in C57BL/6 blastocysts injected with BTBR6 or BTBR 1.3 ES cells

Cell line	Day recovered	Number transferred	Number implanted	Number resorptions	Number grossly retarded	Number correct embryonic stage	Number normal embryos	Number posteriorly abnormal embryos	Number normal chimaeric (% normal)	Number abnormal chimaeric (% abnormal)
BTBR 6 (T/T)	8.5, 9.25 and 9.5	76	66 86.8%	19 28.8%	3* 4.5%	44 66.7%	34 72.3%	10 22.3%	1† 2.9%	10 100%
BTBR 1.3 (T/+)	9.25	21	14 66.7%	6 42.9%	0	8 67.9%	8 100%	0	5 62.5%	0

About 10 BTBR 6 or BTBR 1.3 ES cells, which are both homozygous for the *Gpi 1^b* allele of GPI, were injected into C57BL/6 blastocysts, which are *Gpi 1^b/Gpi 1^b*. Injected embryos were recovered either 6 or 7 days after transfer to MF1 pseudopregnant recipients. Only embryos transferred to recipients which subsequently became pregnant are included. Embryos were dissected free of the decidua and Reichert's membrane was removed. The visceral yolk sac was reflected and dissected away from the embryo to give a clear view of the morphology of the embryo and the allantois. After inspection for abnormalities the allantois and embryonic regions of 8.5-day embryos were analysed separately for relative host and donor GPI contributions. The embryonic portion of the six abnormal BTBR 6 $\leftrightarrow$ +/+ 9.25-day and 9.5-day chimaeras was fixed in Bouin's fluid and processed for routine histology. The allantois and/or amnion were removed for GPI analysis. The single chimaera designated as normal (asterisk) was only 5–10% chimaeric in the embryo and there was no evidence of an ES cell contribution to the allantois. Chimaerism was detected by GPI electrophoresis. The different tissue fractions from each embryo were placed separately in microdrops of PBS under paraffin oil. Samples were frozen and thawed at least twice before loading on to Titan III plates (Helena Laboratories). Electrophoresis and staining for GPI activity followed previously described methods^{21,22} except that the concentration of the Tris-glycine buffer was reduced to 0.025%.

* One grossly retarded embryo was chimaeric.

† 10% chimaerism in the embryo only.

the sequence of the gene nor its expression pattern define its developmental function^{7,9}. To study the cell autonomy of the *T* mutation we have isolated and genetically characterized embryonic stem cell lines and studied their behaviour in chimaeras. *T*/+ embryonic stem cells form normal chimaeras, whereas *T*/T $\leftrightarrow$ +/+ chimaeras mimic the *T*/T mutant phenotype. The results indicate that the *T* gene acts cell autonomously in the primitive streak and notochord but may activate a signalling pathway involved in the specification of other mesodermal tissues.

Preimplantation *T*/T embryos, like many other postimplantation embryonic recessive lethals, are morphologically indistinguishable from wild-type embryos. Therefore, we have used homozygous *T*/T embryonic stem (ES) cells to make chimaeras between mutant cells and normal embryos to ascertain whether the mutant *T* phenotype can be rescued in the presence of wild-type cells. ES cell lines were derived by standard pro-

cedure¹⁰ from blastocysts recovered from BTBR Pas *T*/+ $\times$ *T*/+ matings. Five independent ES cell lines were tested by Southern blot analysis for deletion of the *T* gene (Fig. 2). Two (BTBR6 and BTBR10) were homozygous for the deletion. BTBR 1.3 and 7 are heterozygous, and BTBR 4 is wild-type (data not shown). All five cell lines show similar growth profiles *in vitro*.

Table 1 shows the frequency of normal development among midgestation conceptuses derived from blastocysts injected with BTBR 6 or BTBR 1.3. The 10 abnormal embryos from BTBR 6 injections were indistinguishable from their littermates with respect to anterior morphology but were markedly abnormal caudally. The most distinctive feature of 8.5-day abnormal embryos was either a much reduced allantois (Fig. 1e) or an allantois with an uncharacteristic branching morphology (Fig. 1c, d). As in *T*/T mutants, cells from the base of the allantois tended to spread onto adjacent amniotic mesoderm

FIG. 1 Development of 8.5 and 9.25 d.p.c. *T*/T mutants and chimaeras. *a*, 8.5 d.p.c. *T*/T embryo (asterisk) and normal littermate. Scale bar, 250 μ m. The anterior aspect of the *T*/T embryo is normal but the allantois is reduced (arrow) and the primitive streak region more opaque. *b*, 9.5 d.p.c. *T*/T embryo (asterisk) and normal littermate. Scale bar, 550 μ m. All components of the *T*/T embryo anterior to the seventh somite (arrow) appear normal. Caudally the neural tube is kinked and the somites disorganized. The primitive streak region is thickened and only a rudimentary allantois (small arrow) is present. *c*, 8.5 d.p.c. BTBR 6 $\leftrightarrow$ +/+ chimaera 1. The embryo appears normal but there is a distinctive 'bubble' on the end of the allantois (arrow). Scale, 250 μ m. *d*, 8.5 d.p.c. BTBR 6 $\leftrightarrow$ +/+ chimaera 2, with abnormal 'branched' (arrow) allantois. Scale bar, 230 μ m. *e*, 8.5 d.p.c. BTBR 6 $\leftrightarrow$ +/+ chimaera 3 with stunted allantois. Scale bar, 200 μ m. Cells from the allantois are spreading over the amnion (arrow). *f*, 9.25 d.p.c. BTBR 6 $\leftrightarrow$ +/+ chimaera 5. The anterior third is normal but posterior to somite 7; the trunk and caudal region are abnormal. The primitive streak region is enlarged and the allantois (arrow) minimal. Scale bar, 250 μ m. *g*, Transverse section through the caudal region of BTBR 6 $\leftrightarrow$ +/+ chimaera 9. Pyknotic cells are evident in a discrete central region (dotted line) beneath the primitive streak. The darkly staining nuclei on either side of the dorsal aspect of the gut (G) are normal nuclei of blood cells. Scale bar, 30 μ m. *h*, 9.25 d.p.c. 15 somite BTBR 1.3 $\leftrightarrow$ +/+ chimaera 4. The embryo seems entirely normal and the allantois had fused with the chorion. Scale bar, 400 μ m.

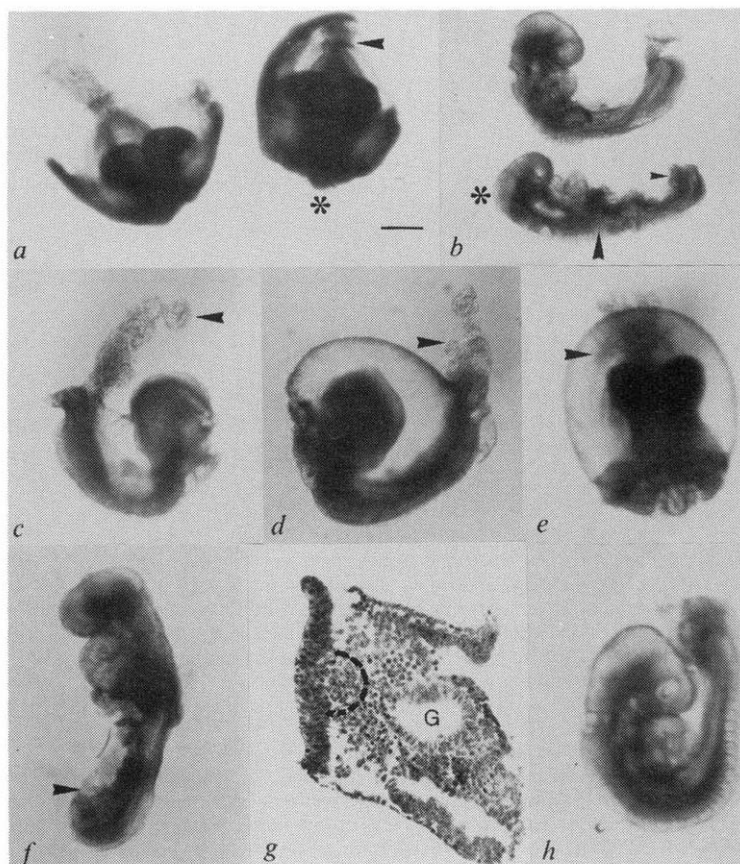


TABLE 2 Extent of chimaerism and developmental characteristics of BTBR6 $\leftrightarrow$ +/+ and BTBR 1.3 $\leftrightarrow$ +/+ mid-gestation chimaeras

	BTBR 6 injections											BTBR 1.3 injections				
Chimaera number	1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5
Day of recovery	8.5	8.5	8.5	8.5	9.25	9.5	9.5	9.5	9.5	9.5	9.25	9.25	9.25	9.25	9.25	9.25
Somite number	7	7	6	2	7+	7	6	7+	7+	7+	15	9	12	15	15	9
Abnormal posterior	+	+	+	++	+++	+++	++	++	++	+++	—	—	—	—	—	—
Allantois morphology	+	++	+++	+++	+++	+++	+	++	++	+++	—	—	—	—	—	—
% GPI-1A in embryo	10	25	60	60	50	ND	ND	ND	ND	ND	10	50	0	10	50	100
% GPI-1A in allantois	10	50	60	ND	90	ND	10	ND	30	80	0	90	15	10	30	100
% GPI-1A in amnion	ND	ND	ND	ND	ND	100	ND	90	ND	80	ND	ND	ND	ND	ND	ND

The relative contribution of the two GPI allozymes was assessed visually after staining and fixation of the Titan III plates. BTBR 6 $\leftrightarrow$ +/+ 9.25 and 9.5-day chimaeras had disorganized somites in addition to seven apparently normal rostral somites (denoted by 7+). Posterior abnormality is classified as follows: +, thickened or asymmetrical primitive streak; ++, kinked neural tube, disrupted somites, thickened primitive streak; +++, neural tube and somites abnormal posterior to the forelimb bud region. Allantoic morphology is classified as follows: +, blebs on the surface of the allantois; ++, blebs and branching morphology; +++, grossly reduced allantois with a tendency to spread onto the amnion. The extent of development in 8.5-day embryos means that usually only primitive streak and allantoic morphology can be accurately scored.

(Fig. 1e) and the primitive streak region seemed unusually dense. At 9.25 or 9.5 days postcoitum (d.p.c.) (Fig. 1f) the somites caudal to the forelimb bud region appeared small and irregular in shape and distribution. Fewer somites were present than would be expected for the stage of cranial development. The neural tube was severely kinked and the primitive streak region clearly disorganized. There was only a minimal allantois. Histological sections of the caudal region of six abnormal 9.25-day and 9.5-day embryos showed that the notochord was either discontinuous or contained pyknotic cells and that dead cells were evident among nascent mesoderm emerging from the primitive streak (Fig. 1g). Dead cells were not prominent in sections of abnormal allantois from chimaeras 4, 6 and 8. Overall, the 10 abnormal embryos seem to mimic the phenotype of *T/T* mutant embryos.

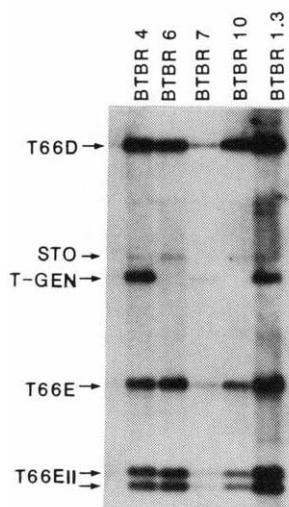


FIG. 2 Southern blot determination of the genotype of BTBR ES cell lines. High molecular weight DNA was extracted from five independent BTBR lines and digested with *TaqI* restriction enzyme. After electrophoresis the DNA was transferred to a nylon filter (Hybond, Amersham) and hybridization to probes carried out in 0.5M sodium phosphate buffer (pH 6.8) containing 7% SDS at 68 °C, before washing in 40 mM sodium phosphate buffer (pH 6.8) containing 1% SDS at 60 °C (ref. 17). The filter was first probed with p19OR10RS, which is a single-copy 1.2-kilobase fragment from the promoter region of the *T* gene (T-GEN). This probe also recognizes the *T* gene in the STO feeder cells, which is polymorphic between the 129/Sv genetic background of BTBR/Pas and SIM, from which STO cells are derived. Consequently, the lower band T-GEN is specific to the ES cells whereas the upper band (STO) is specific to the remaining feeder cells. The blot was rehybridized with p66MRT (refs 18, 19), which recognizes flanking sequences on the centromeric (T66E and T66EII) and telomeric (66D) sides of the *T* gene. Both probes were 32 P-labelled by random primer extension²⁰. BTBR 6 and 10 are homozygous for the *T* deletion. BTBR 4, 7 and 1.3 contain at least one copy of the wild-type *T* allele, and subsequent analysis, probing for a restriction length polymorphism diagnostic for the deletion junction region, shows that BTBR 4 is +/+ whereas BTBR 7 and 1.3 are heterozygous (data not shown).

Glucose phosphate isomerase (GPI) analysis (Tables 1 and 2) revealed that all 10 abnormal embryos were chimaeric. One apparently normal embryo proved to be slightly chimaeric (less than 10%) but the allantois contained only host cells. By contrast, all eight conceptuses recovered from BTBR 1.3 injections were morphologically normal. Of these, five were chimaeric and one showed no trace of the host blastocyst GPI allozyme in the embryo or allantois (Table 2). Viable heterozygous *Brachyury* chimaeras have been described before in the chimaeric rescue of cells carrying a maternally derived *T^{hp}* deletion¹¹.

The replication of the *T/T* phenotype in BTBR 6 $\leftrightarrow$ +/+ chimaeras means that these null ES cells provide an ideal vehicle for attempting the molecular manipulation of mesoderm formation, using all or part of the *T* gene, or genes acting downstream of it. In addition, introduction of a single cell marker, such as *lacZ*, into these cell lines will allow a detailed analysis of the cellular basis of aberrant development.

The principal mesodermal tissues affected in *T/T* mutants are those situated in the midline, which do not form until much of the cranial, cardiac and extraembryonic mesoderm has already left the streak¹². The discontinuous notochord and presence of pyknotic cells in the notochord and primitive streak of 9.25-day and 9.5-day BTBR 6 $\leftrightarrow$ +/+ chimaeras, indicate that the *T* gene acts cell autonomously in these tissues. But, the consistent disruption of the allantois, which does not express the *T* gene⁸, and the spreading of allantoic cells onto the amnion (Fig. 1e), suggests that the *T* product acts during the specification rather than differentiation of this tissue. There is no apparent selection against *T/T* cells contributing to the allantois (Table 2) and there is no evidence for elevated cell death in this region, contrary to what would be expected if *T* acted cell autonomously in primitive streak progenitors of the allantois. Therefore, we would postulate that the *T* gene product is required to maintain the notochord and its precursors in the primitive streak, but that it also has a critical role upstream of a signalling pathway required to specify other mesodermal derivatives, such as the allantois, which themselves do not express the *T* gene. The fact that this signalling role cannot be rescued in chimaeras indicates that important threshold effects operate. Evidence from the graded anteroposterior severity of defects seen in different *T* alleles^{13–16} argues that mesoderm formation has an increased demand for the *T* product as development proceeds. Thus, a temporal gradient in the level of *T* expression in the streak could be responsible for specifying different mesodermal derivatives and ensuring a progressively more posterior address for nascent mesoderm. □

Received 24 May; accepted 22 July 1991.

1. Dobrovolskaia-Zavadskaja, N. *C.R. Séanc. Soc. Biol.* **97**, 114–116 (1927).
2. Gluecksohn-Schonheimer, S. *Proc. natn. Acad. Sci. U.S.A.* **30**, 134–140 (1944).
3. Chesley, P. *J. exp. Zool.* **70**, 429–459 (1935).
4. Gruneberg, H. *J. Embryol. exp. Morph.* **6**, 424–443 (1958).
5. Yanagisawa, K. O., Fujimoto, H. & Urushihara, H. *Dev. Biol.* **87**, 242–248 (1981).
6. Fujimoto, H. & Yanagisawa, K. O. *Differentiation* **25**, 44–47 (1983).
7. Herrmann, B. G., Labeit, S., Poustka, A., King, T. R. & Lehrach, H. *Nature* **343**, 617–622 (1990).
8. Wilkinson, D., Bhatt, S. & Herrmann, B. G. *Nature* **343**, 657–659 (1990).

9. Willison, K. *Trends Genet.* **6**, 104-105 (1990).
10. Robertson, E. J. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (ed. Robertson, E. J.) 71-112 (IRL, Oxford, 1987).
11. Bennett, D. *Nature* **272**, 539 (1978).
12. Tam, P. P. L. & Beddington, R. S. P. *Development* **99**, 109-126 (1987).
13. Lyon, M. F. *J. Hered.* **50**, 140-142 (1959).
14. Searle, A. G. *Genet. Res.* **7**, 86-95 (1966).
15. MacMurray, A. & Shin, H.-S. *Genetics* **120**, 545-550 (1988).
16. Yanagisawa, K. O. *Jap. J. hum. Genet.* **65**, 287-297 (1990).
17. Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991 (1984).
18. Herrmann, B. *et al. Cell* **44**, 469-476 (1986).
19. Herrmann, B., Barlow, D. P. & Lehrach, H. *Cell* **48**, 813-825 (1987).
20. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6 (1983).
21. McLaren, A. & Buehr, M. *Genet. Res. Camb.* **37**, 305-309 (1981).
22. Eicher, E. M. & Washburn, L. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 946-950 (1978).

ACKNOWLEDGEMENTS. We thank A. Authaler, D. Weigel and L. Manson for technical assistance, F. Johnston for preparation of photographs and A. Smith for discussion. P.R. is supported by the Imperial Cancer Research Fund.

Drosophila genes *Posterior Sex Combs* and *Suppressor two of zeste* encode proteins with homology to the murine *bmi-1* oncogene

Brian P. Brunk*, Elisabeth C. Martin & Paul N. Adler

Biology Department, Molecular Biology Institute and Cancer Center, University of Virginia, Charlottesville, Virginia 22901, USA

* Present address: Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111, USA

THE *Polycomb* group (*Pc-G*) genes are needed to maintain expression patterns of the homeotic selector genes of the *Antennapedia* (*Antp-C*) and *bithorax* (*bx-C*) complexes, and hence for the maintenance of segmental determination¹⁻³. We report the

predicted protein sequence of the *Pc-G* gene *Posterior Sex Combs* (*Psc*), and of the neighbouring and related gene *Suppressor two of zeste* (*Su(z)2*). Both genes encode large proteins that contain a 200 amino-acid domain identical over 37.4% that is also conserved in the murine oncogene *bmi-1* (refs 4, 5). At the amino terminus of this domain is a cysteine-rich sequence that has been proposed as a novel type of zinc finger⁶.

It is clear that some *Pc-G* genes are pleiotropic regulatory genes. Mutations in *Psc* can cause segmental transformations⁷, ventral toward dorsal transformations⁸, abnormal axon outgrowth in the embryonic central nervous system (CNS) (D. Smouse, N. Perrimon and C.-t. Wu, personal communication) and suppression of the *z¹-w* interaction⁹. Conceptual translation of a 6,031-base pair (bp) *Psc* complementary DNA clone^{10,11} (detailed characterization of *Psc* will be presented elsewhere) yields a single long ORF encoding 1,603 amino acids (EMBL accession number X59275); conceptual translation of 6.3-kilobases (kb) of *Su(z)2* cDNA (refs 10, 11) yields a single long ORF encoding 1,365 amino acids (EMBL accession number X56799). The large predicted *Psc* and *Su(z)2* proteins share a number of features including PEST sequences¹², many potential

FIG. 1 Graphical demonstration of the dual nature of the relatedness of *Psc* and *Su(z)2*. *a*, PLFASTA comparison of the *Psc* and *Su(z)2* proteins¹⁴. The *Psc* protein is represented along the horizontal axis and the *Su(z)2* protein is represented along the vertical axis. The shaded region denotes the HR and is intended to highlight the uniqueness of this region relative to the rest of the protein. Only one alignment is found in the HR, whereas many are seen in the SAACR. The line type indicates the FASTA score for each of the alignments. The alignment with the highest score is the one alignment that encompasses the HR. *b*, Flexibility plots of the *Psc* and *Su(z)2* proteins obtained using the IBI Pustell software (higher values indicate greater flexibility). The plots are aligned so that the HRs are lined up with the HR of *Psc* in *a*. For both proteins the HR is markedly less flexible than the rest of the protein (SAACR), another indication of the unique nature of the HR. Secondary structure prediction programs predict very similar amounts of structure motifs for both SAACRs. For example, the Garnier program from PCGENE (Intelligenetics) predicts 31.5% of the *Psc* SAACR and 32.2% of the *Su(z)2* SAACR to be in α helix, 41.1% of *Psc* and 39.8% of *Su(z)2* to be in an extended configuration (β sheet), 10.8% of *Psc* and 12.6% of *Su(z)2* to be in reverse turns and 16.4% of *Psc* and 15.2% of *Su(z)2* to be in a random coil. Note that it is possible for two proteins to have similar amino-acid compositions but different predicted structural properties if the distribution of particular amino acids is different for the two proteins.

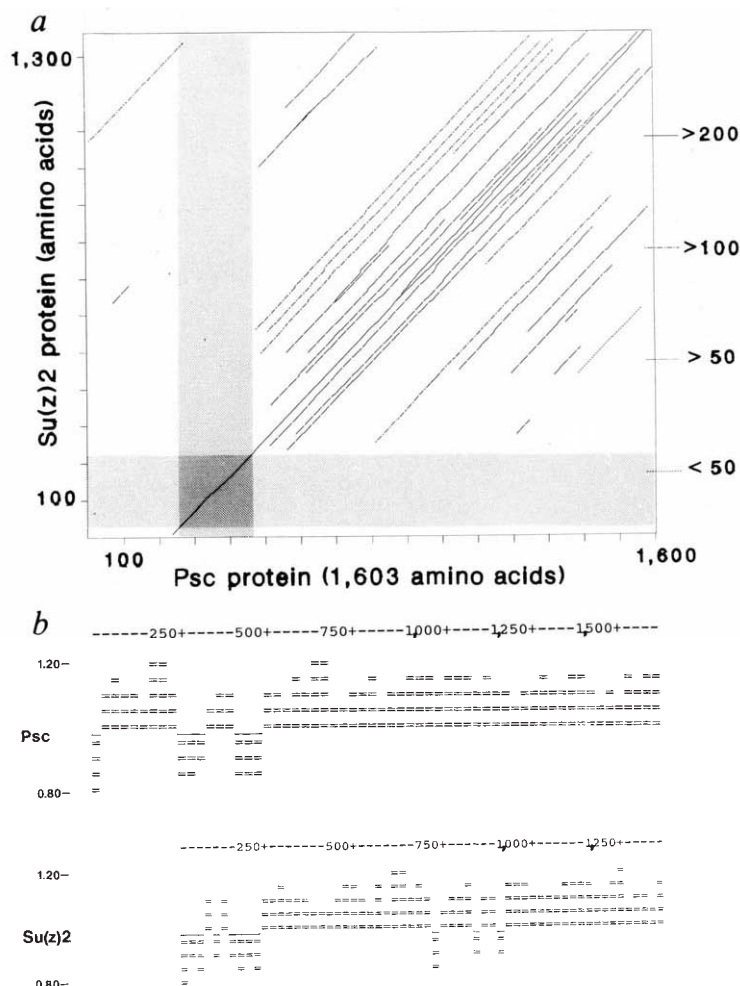


FIG. 2. Sequence similarity between *Psc*, *Su(z)2* and *bmi-1* proteins (single-letter amino-acid code). The FASTA alignment of the homology regions of *Su(z)2* and *bmi-1* to that of *Psc* is shown. A short stretch of amino acids preceding and following the HR is also shown. The HR, as defined by comparing *Psc* and *Su(z)2*, is demarcated by arrows and the putative zinc-finger region is in bold type. *Psc* and *Su(z)2* are 37.4% identical in the HR, *Psc* and *bmi-1* 43% identical and *Su(z)2* and *bmi-1* 32.1% identical.

phosphorylation and glycosylation sites, potential nuclear localization sequences¹³, runs of a single amino acid, an abundance of proline and serine residues and a number of regions of high positive or negative charge. Sequence comparison of *Psc* and *Su(z)2* yields one alignment that extends the entire length of *Su(z)2*, as well as many shorter ones¹⁴ (Fig. 1a). Only one region of alignment shows actual 'sequence' identity. This homology region (HR) consists of a 200 amino-acid portion that is 37.4% identical (Fig. 2) beginning 30 amino acids from the N-terminal end of *Su(z)2* and 264 amino acids from the N-terminal end of *Psc*. Sequence-shuffling analysis¹⁴ shows that in the remainder of the two proteins (85% of *Su(z)2* and 87% of *Psc*) the alignments are not due to 'sequence identity', but rather to the very similar amino acid content in this region (SAACR) (Table 1). The two SAACRs are predicted by a variety of programs to have similar amounts of different secondary structures, and the HRs are predicted to be markedly less flexible than the SAACRs (Fig. 1b).

The murine *bmi-1* oncogene^{4,5} is the only protein we identified as containing a long region of sequence similarity to *Psc* and/or *Su(z)2*. Indeed *Psc* has been cloned independently by its sequence similarity to *bmi-1* (ref. 15). Remarkably, the region of *Psc* (and *Su(z)2*) that is similar to *bmi-1* corresponds almost

```

Homology region sequence similarity
+
DVRQFHDLLITCRLCRGYMIDPTTVDYCYHTYCRSCILKHLRLRAVYCECKASGGKEINE
. . . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
LLTAVNPHIIICHLCCQGYLINATTIVECLHSFCHSCLINHLRKRKFCPCRCCEMVINN--AK
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
KITELNPHLMCVLCGGYFIDATTIIECLHSFCKTICIVRYLETSKYCYPICDVQVHKTRPL
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

DTRLRSLIYKLVPGLYQRE--CKELADFKEQHD-----LVDEQTTDEPE---FFTTELI
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PTLQAIIVYKLVPGLYERELMRKRAFVKDRPE--EAALATPQQRGD--DTEHLIFSPSDDM
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
KTLQDIIVYKLVPGLFKNEMKRRRRDYAAHPSADAANGSNEDRGEVADEEKRIITDDEII
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

-HPAML--HQCGPGEVPPTI---YLQCAAGLPVELLKRFLCSKYNIETDNLGVEVEVTY
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
AELGEL--KTDSEPELVDTLRLPRYLQCPAMCRVSHLKKFVYDKFEIDAQRFSIDIMYKV
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
FDQRS�DRKVNKEPKPEEVNDKRYLRCPAAMTVMHLRKLRLSRKMDI--PNTFQIDVMYEE
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

TNFTLMDVAYCYNWSRESPMFACYRILLYDNEQTKNDENNLS
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OYYTLMEDIAITYTWKRDAPMRFYYRV--YESQPQLVKPAPRR
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OYYTLMEDIAIYITWRRNGPLPLKYRV--RPTCKRMKMSHORD
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

exactly to the HR (Fig. 2). Furthermore, the intron-exon structure of *Su(z)2* reflects the sequence relationship between *Su(z)2* and *Psc*, as the HR of *Su(z)2* is encoded by three small exons separated by split codons, whereas the SAACR region of *Su(z)2* is encoded by one large exon^{16,17}. These data suggest that the HR encompasses a new type of protein domain. At the N-terminus of the HR is a consensus sequence for a proposed new family of zinc-finger proteins⁶ (Fig. 3). Many of the members of this group are likely to be DNA-binding proteins, and for VMW110, a transcription factor from HSV, this cysteine-rich region is required for activity¹⁸.

It seems likely that the equivalent of *Pc*-G genes will be found in mammals. *Psc* is more closely related to *bmi-1* than to *Su(z)2*, consistent with *bmi-1* being a mammalian *Pc*-G gene, and the murine equivalent of *Psc*; however, the relatively low degree of identity (43% in the HR) and the much smaller size of *bmi-1* make this uncertain. Consistent with the sequence similarities between *Psc*, *Su(z)2* and *bmi-1* reflecting functional similarities, aberrant expression of any of these genes can result in abnormal differentiation (refs 6, 7, 11 and E. J. Sharp and P.N.A., unpublished results). It is intriguing, but not surprising, that a *Pc* group gene is related to an oncogene. In the most general sense, the *Pc* group genes are required to maintain the determined

Su(z)2	I T C R L C R G Y M I D P T T V D - - - - - Y Q Y H T Y C R S C I L K H L L R A V - - - - - Y C F E C K
Psc	I I C H L C Q G Y I N A T T V - - - - - E C L H S C F G S C L I N H L R K E R - - - - - F C P R E E
BMI-1	L M C V L C G G Y I D A T T I I - - - - - E C L H S C F K T C I V R Y L E T S K - - - - - Y C H I O D
RAD18	L R C H I C K D F I K V P V L T - - - - - P G G H T F C S L C R T H L N N Q P - - - - - N C P L L L
RAG1	I S C Q I C E H I A D P V E T - - - - - S C K H L F R C I L R C L K V M G S - - - - - Y C F S C R
HUMRFP	T T C P V C L Q Y F A E P M M L - - - - - D C G H N I C C A C L A R C W G T A E T N V - - - - - S C F Q C R
HUMROS	V T C P I C L D P F V E P V S I - - - - - E C G H S C F Q E C I S Q V G K G G G S - - - - - V C P V C R
MUSRPT	V T C P I C L D F E K P V S A - - - - - D C N H S C F R A C I T L N Y E S N R N T D G K G N C P V C R
WZBE61	N T C T I C M S T V S D L G K T M - - - - - P C L H D F C F V C I R A W T S T S V - - - - - Q C F L C R
Vmw110	D V C A V C T D E I A P H L R C D T F - - - - - P C M H R F C I P C M K T W M Q L R N - - - - - T C P L C N
TLR	G M C A V C R E P J A G A V E L L - - - - - P C M H V E G T A C V Q R W - - - - - R C S C Q
RING1	L M C P I C L D M I K N T M T T K - - - - - E C L H R F C S D C I V A L R S G N K - - - - - E C P T C R
CG30	L Q G N I C F S V A E I K N Y L F Q P I D R L T I I P V L E L D T C K H Q I C S M C R K I R K R K K V - - - - - P C F L C R
consensus	C Ψ C Ψ C H Ψ C CΨ CP C

FIG. 3 The putative *Psc*, *Su(z)2* zinc-finger homology sequence. The sequences of the 11 putative zinc-fingers from proteins identified in the databases are shown aligned with the putative zinc-finger region of *Su(z)2* and *Psc*. A derived consensus sequence is shown below. Mutations in RAD18 result in increased sensitivity to agents that damage DNA¹⁹. RAG1 has been implicated in V(D)J recombination in the immunoglobulin family²⁰, and the MURPT protein negatively regulates expression of the interleukin-2 receptor gene and genes regulated by the long-terminal repeat of human immunoglobulin.

deficiency virus²¹. The HUMROS protein is a component of the Ro antigen which is a ribonucleoprotein²². TLR is a trypanosome gene identified by its location near an expressed copy of a trypanosome surface antigen²³. WZBE61 was identified as an open reading frame in the varicella zoster virus²⁴, and HUMRFP was identified as part of an oncogene that resulted from a protein fusion²⁵. RING1 is a gene located near the DP region of the major human histocompatibility locus⁶. Finally, CG30 is a gene from baculovirus²⁶.

TABLE 1 Amino-acid composition of the SAACRs

Code	<i>Su(z)2</i>	<i>Psc</i>	Swiss Prot
Ala	11.1	10.9	7.7
Arg	3.9	2.9	5.3
Asn	7.4 ⁺	6.0	4.4
Asp	2.5 ⁻	2.0 ⁻	5.2
Cys	0.7 ⁻	0.4 ⁻	1.8
Gln	6.3	5.5	4.1
Glu	4.1	4.2	6.3
Gly	4.3	5.6	7.1
His	1.6	1.2	2.3
Ile	2.9	2.9	5.4
Leu	7.2	6.0	9.1
Lys	7.3	8.8	5.9
Met	2.1	2.2	2.3
Phe	1.1 ⁻	1.0 ⁻	4.0
Pro	10.0 ⁺	12.3 ⁺⁺	5.1
Ser	15.0 ⁺⁺	12.8 ⁺	7.1
Thr	6.9	8.3	5.9
Trp	0.0 ⁻	0.0 ⁻	1.3
Tyr	1.1 ⁻	1.2 ⁻	3.2
Val	3.6	4.7	6.5

The amino-acid composition of the SAACRs of *Psc* and *Su(z)2* are shown. For both SAACRs every amino acid deviated in the same direction (higher or lower) from the frequency found in the Swiss Prot database (this data was obtained from the PCGENE program (Intelligenetics)). Values for the amino-acid composition of the SAACR are given as percentages, a - or a + after a value indicates that the value differs from the average database value (Swiss Prot release 16) by greater than 50%. ++, The value is greater than two-fold higher; --, the value is less than 25% of the average value.

state. Cancer cells typically lose some of their differentiated properties; thus it seems reasonable that mutations that interfere with the genetic circuitry required for the maintenance of embryonic determination could lead to oncogenesis.

Note added in proof: Immunostaining of polytene chromosomes shows that the *Psc* protein is a locus specific chromosomal protein present at the location of the *Antp-C*, *bx-C* and about 40 other locations (E.C.M. and P.N.A., unpublished; L. Rastelli and V. Pirrotta, personal communication). □

Received 17 June; accepted 6 August 1991.

- Struhl, G. & Akam, M. *EMBO J.* **4**, 3259-3264 (1985).
- Wedeen, C., Harding, K. & Levine, M. *Cell* **44**, 739-748 (1986).
- Paro, R. *Trends Genet.* **5**, 131-135 (1990).
- van Lohuizen, M. et al. *Cell* **65**, 737-752 (1991).
- Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P. & Adams, J. M. *Cell* **65**, 753-763 (1991).
- Freemont, P. S., Hanson, I. M. & Trowsdale, J. *Cell* **64**, 483-484 (1991).
- Jurgens, G. *Nature* **316**, 153-155 (1985).
- Adler, P. N., Martin, E. C., Charlton, J. C. & Jones, K. *Dev. Genet.* (in the press).
- Wu, C. T., Jones, R. S., Lasko, P. F. & Gelbart, W. M. *Molec. Gen. Genet.* **218**, 559-564 (1989).
- Brunk, B. P. & Adler, P. N. *Genetics* **124**, 145-156 (1990).
- Brunk, B. P., Martin, E. C. & Adler, P. N. *Genetics* **128**, 119-132 (1991).
- Rogers, S., Wells, R. & Rechsteiner, M. *Science* **234**, 364-369 (1986).
- Chelsky, D., Ralph, R. & Jonak, G. *Molec. cell. Biol.* **6**, 2487-2492 (1989).
- Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444-2448 (1988).
- van Lohuizen, M., Frasch, M., Wientjes, E. & Berns, A. *Nature* **353**, 353-355 (1991).
- Brunk, B. P. & Adler, P. N. *Nucleic Acids Res.* **19**, 3149 (1991).
- Brunk, B. P. thesis, Univ. Virginia (1989).
- Everett, R. D. *J. molec. Biol.* **202**, 87-96 (1988).
- Jones, J. S., Weber, S. & Prakash, L. *Nucleic Acids Res.* **16**, 7119-7131 (1988).
- Schatz, D. G., Oettinger, M. A. & Baltimore, D. *Cell* **59**, 1035-1048 (1989).
- Patarca, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2733-2737 (1988).
- Itoh, K., Itoh, Y. & Frank, M. B. *J. clin. Invest.* **87**, 177-186 (1991).
- Smiley, B. L., Stadnyk, A. W., Myler, P. J. & Stuart, K. *Molec. cell. Biol.* **10**, 6436-6444 (1990).
- Davison, A. J. & Scott, J. E. *J. gen. Virol.* **67**, 1759-1786 (1987).
- Takahashi, M., Inaguma, Y., Hiai, H. & Hirose, F. *Molec. cell. Biol.* **8**, 1853-1856 (1988).
- Theim, S. & Miller, L. K. *J. Virol.* **63**, 4489-4497 (1989).

ACKNOWLEDGEMENTS We thank M. van Lohuizen for letting us know about his cloning of a *Drosophila* homologue of *bmi-1*, and C. Vinson, C. Emerson, W. Pearson, A. Beyer, J. Charlton, S. Conover and M. Allison. This work was supported by the NSF, B.P.B. and E.C.M. were supported by a NIH predoctoral training grant in developmental biology.

Sequence similarity between the mammalian *bmi-1* proto-oncogene and the *Drosophila* regulatory genes *Psc* and *Su(z)2*

Maarten van Lohuizen*, Manfred Frasch†‡, Ellen Wientjes* & Anton Berns*

* Division of Molecular Genetics, The Netherlands Cancer Institute and Department of Biochemistry, the University of Amsterdam, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

† Max-Planck Institut für Entwicklungsbiologie, Abteilung Genetik, D-7400 Tübingen, Germany

THE *bmi-1* proto-oncogene can be activated by Moloney murine leukaemia proviral insertions in *Eμ-myc* transgenic mice^{1,2}. It encodes a highly conserved nuclear protein of 324 amino acids which belongs to a family of proteins containing a putative new zinc-finger. Another closely related member of this family is the mouse protein Mel-18 (ref. 3). Here we report on the cloning and characterization of a homologous gene (*D-bmi*) from *Drosophila melanogaster*. Our analysis indicates that distinct domains of the mouse *Bmi-1* protein, including the putative zinc-finger motif, are highly conserved within the much larger *D-Bmi* protein. Chromosomal localization and sequence comparison reveal that *D-bmi* is identical to *Posterior Sex Combs* (*Psc*)⁴⁻⁶ and indicate that the conserved domains between mouse *bmi* and *Psc* are also conserved within *Suppressor-2 of Zeste* (*Su(z)2*)⁶⁻⁸.

We used a mouse *bmi-1* complementary DNA probe to search for a *Drosophila* homologue. This probe detected specific bands on a *Drosophila* Southern blot when used under low-stringency hybridization conditions¹ (Fig. 1a). From a *Drosophila* genomic library screened under identical conditions, we obtained several overlapping clones, and fine-mapping of one of these indicated that the homology was contained within an 800-base pair *SalI*-*HindIII* fragment (probe A). Probe A detected a transcript of 6.6 kilobases (KB) on northern blots of RNA from different developmental stages (Fig. 1b). Screening of an embryo cDNA library⁹ gave several almost full-length clones. Three of these clones were partially sequenced and the derived amino-acid sequence aligned with the mouse *Bmi-1* sequence (Fig. 2).

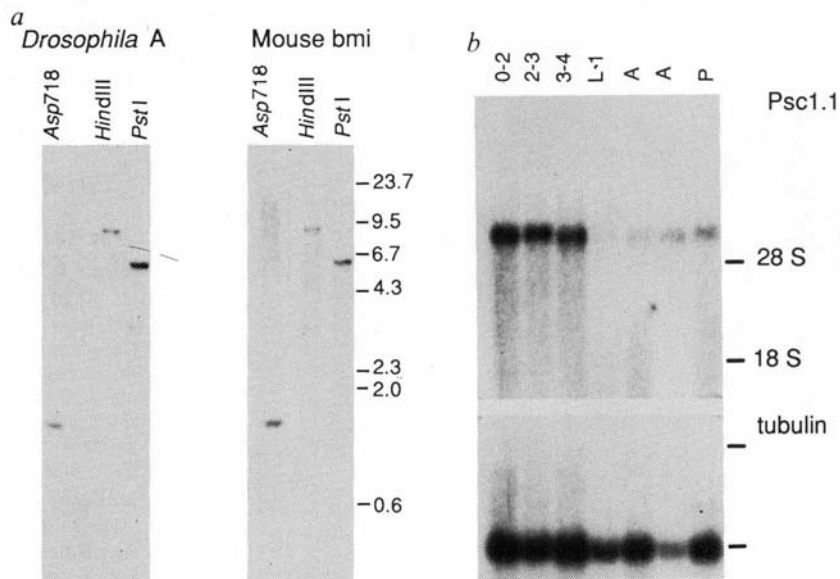
Hybridization of polytene chromosomes with a *D-bmi* probe showed that the *D-bmi* gene is localized at 49E on chromosome 2R (data not shown). This region contains several genes identified by specific mutations, affecting the regulation of homeotic genes and other developmental processes^{4,5,10}. Two of these genes, *Posterior Sex Combs* and *Suppressor-2 of Zeste* have been cloned by P-element tagging^{10,11}. Comparison of *D-bmi* with *Psc* and *SU(z)2* sequences revealed that *D-bmi* is identical to *PSC* (ref. 6). The sequences conserved between mouse *bmi-1* and *Psc* are also conserved within *SU(z)2*. Conservation is particularly strong in the putative zinc-finger domain, extending beyond the previously derived consensus sequence^{1,2}. In addition, the second part of a putative helix-turn-helix domain¹ is identical between *Psc* and mouse *Bmi* proteins and is also highly conserved within *Su(z)2* (see alignment in Fig. 2). The mouse *Bmi-1* protein has 41% identity and 61% similarity with *Psc*, and 34% identity and 54% similarity with *Su(z)2*, respectively (data not shown). Although the amino-acid sequence of Mel-18 is identical to mouse *Bmi-1* in the region of homology with *Psc* and *Su(z)2*, it is more distantly related at the DNA level owing to third base substitutions (see legend to Fig. 1).

Loss-of-function mutants in *Psc* are embryonic lethals displaying head defects and partial transformation of several seg-

‡ Present address: Mount Sinai Medical Center, 1 Gustave L. Levy Place, New York, New York 10029, USA.

FIG. 1 Cross-hybridization of mouse *bmi-1* orf sequences with *Drosophila* DNA fragments, and detection of D-*bmi* transcripts. **a**, Right panel: Southern blot of *Drosophila* DNA digested with *Asp*718, *Hind*III or *Pst*I was hybridized under reduced stringency with a mouse *bmi-1* polymerase chain reaction fragment containing only open reading frame sequences¹. Numbers on the right indicate size markers (kb). Left panel: Duplo-Southern blot, hybridized under normal stringency with D-*bmi* probe A, which contains part of the conserved sequences distributed over two exons. **b**, Northern blot containing *Drosophila* RNA from several developmental stages hybridized with a D-*bmi* 1.1 cDNA probe, and re-hybridized with an α -tubulin probe to control for the amount of RNA loaded. The same 6.6-kb D-*bmi* transcript was also detectable in 4–24 h embryos (not shown). Numbers above first three lanes refer to hours of embryonic development. First instar larvae (L1); pupae (P); adult (A). Migration of ribosomal RNA markers (18S and 28S) is shown on the right.

METHODS. Low-stringency hybridization was in 35% formamide at 35 °C, as described in ref. 1. Final wash was in 2 × SSC at 55 °C. Genomic D-*bmi*/*Psc* clones were isolated from a lambda-FIX library (prepared in the laboratory of G. M. Rubin) from *Drosophila* w, th, st, cu, sr, e^s, ca DNA. cDNA clones were isolated from an 8–12-h library (made by N. H. Brown) in the pNB40 expression vector⁹. Sequencing: D-*bmi* cDNA and genomic clones were sequenced on both strands using double-stranded templates and a T7-sequencing kit (Pharmacia) with synthetic oligonucleotide primers. Specificity: although the conservation at the amino-acid level between *Su(z)2* and



Psc/D-*bmi* is high, no cross-hybridization between these genes or their transcripts is detected using the probes and conditions described here. This is probably due to third base substitutions (not shown). Although the mouse Mel-18 protein sequence is identical to mouse Bmi-1 in the conserved region, Mel-18 probes do not recognize *Psc* or *Su(z)2* sequences under our hybridization conditions because of third base substitutions (data not shown). Overall, mouse Bmi-1 and Mel-18 proteins are 70% identical^{1,3}.

ments into more posterior ones, as well as partial ventral-to-dorsal transformations^{4,5,8}. This phenotype is characteristic of genes belonging to the *polycomb* (*Pc*)-group, which are involved in maintaining the segment-specific repression of homeotic selector genes^{4,12–14}. The prototype of this group, Polycomb, has been shown to bind specifically to both homeotic gene loci and to other *Pc*-group genes on polytene chromosomes, probably acquiring sequence specificity through interaction with other proteins^{12,15}. The phenotype of double mutants indicates that *Pc*-group genes act synergistically. Genetic analysis indicates that maternal rescue is responsible for the relatively weak

phenotype of *Psc* and other members of this group^{5,8}. Therefore the enhanced phenotype in double mutants may result from a strong reduction of the maternal rescue. The synergy may be explained by the occurrence of cooperative protein-protein interactions between individual members of the *Pc*-group⁵. Posterior-directed transformations of *Pc*-group mutants, the association of Polycomb with homeotic loci and the conservation of a domain (chromo-domain) between Polycomb and *Su(var)205*, a heterochromatin-associated protein¹⁶, has led to a model in which *Pc*-group members facilitate or regulate the formation of a condensed chromatin configuration along the chromosome,

FIG. 2 Triple alignment of the reduced *Psc*/D-*Bmi* protein sequence with mouse Bmi-1 and *Su(z)2* protein sequences. Only the conserved domain is shown. For complete sequences of *Psc* and *Su(z)2*, see the accompanying paper⁶. Identical amino acids are marked with an asterisk; conserved residues with dots. The conserved putative zinc-finger is stippled, and the conserved helix-turn-helix is marked by an over-bar. The mouse Bmi-1 protein has 41% identity and 61% similarity with *Psc*, and 34% and 54% with *Su(z)2*, respectively. The *Su(z)2* protein sequence was provided by P. N. Adler⁶. Triple alignments were achieved using the CLUSTAL program²².

Bmi-1	1	MH-----RTRIKITELNPHLMCVLCGGYFIDATTIECLHSFCKTCI
<i>Psc</i>	231	AAVAASSTATTTSDLATTSRPRVLLTAVNPHIICHLCQGYLINATTIVECLHSFCHSCL
<i>Su(z)2</i>	1	MHLQNTHTNTMESNAAMDATSPASRDVRFHDLITCRLCRGYMDPTTVDYCYHTYCRSCI
		 * * * * *
Bmi-1	44	VRYLETSKYCPICDVQVHKTRPLLNIKSDTKLQDIVYKLVPGLFKNEMKRRRDFYAAHPS
<i>Psc</i>	292	INHLLRERFCPCRCMVINNAKP--NIKSDTTLQAIYKLVPGLYERELMRKRAFVKDRPE
<i>Su(z)2</i>	61	LKHLRAVYCFCEKASGGKEINELNLSDDTLRLSLIKLVPGLYQRECKELADFKQHDH
		 * * * * * * * * *
Bmi-1	104	ADAANGSNEDRGEVADEEKRIITDDEIISLSIEFFDQSRDLDRVNKEKPKKEEVNDKRYLR
<i>Psc</i>	349	-EAALATPEQRGD--DTEHLIFSPSDDMSLSLEYAELGEL--KTDSEPELVDTLRPRYLQ
<i>Su(z)2</i>	121	VD-----EQTTDEPEFFTTTELISLSLEY-----HPAMLHQCGPGEVPTTIYLO
		 * * * * *
Bmi-1	164	CPAAMTMVHLKFLRSKMDI-PNT--FQIDVMYEEEP--LKDYITLMDIAYIYTWRRNGP
<i>Psc</i>	405	CPAMCRVSHLKFVYDKFEIDAQR--FSIDIMYKVKITVLLDYITLMDIAYIYTWKRDP
<i>Su(z)2</i>	165	CAAGLPVELLKRFLCSKYNITDNLGLVEVEVTKDE--VLPTNFTLMDVAYCYNWRES
		* * * * * * * * * * * *
Bmi-1	219	LPLKYRVRPTCKRMKMSHQDGLT-----NAGELESDDSGSDKANSAGGVPTSSSCLPS
<i>Psc</i>	463	MRFYRYVYESPQPLVKPAPRRVLPLKLEKQERENQEQLAVEVASSKVEPVSLAEDQKAE
<i>Su(z)2</i>	224	MAFCYRI-----LLYDNEQTKNDENNLS--RINQDIEPEHSVRRSKSA
		 * *
Bmi-1	273	PSTP-VQSPHPQFPHISSMTMNGTSNPSAN-----HQSSFASRPR-KS
<i>Psc</i>	523	ASIK-VEGEESTREIVKEVIKDVAATPTTETLKLVINRNMLDKREKSHSPQLSSKSSSKS
<i>Su(z)2</i>	265	KSVTFAEDLESDSGSPRSKVRCKTPPKVSPSSKNKRLTSSKREAEFSPVSNFKSLRS
		* * *
Bmi-1	314	SLNGSSATSSG
<i>Psc</i>	581	SPCTPVSSPSE
<i>Su(z)2</i>	325	N-----DMRY

thereby causing permanent and heritable repression of homeotic genes^{12,17,18}.

Interestingly, several mutant alleles of both *Psc* and *Su(z)2* display a similar dominant gain-of-function phenotype resulting in the suppression of the effect of *Zeste*, a gene coding for a nuclear protein involved in *trans*-sensing mechanisms, which is associated with somatic pairing of chromosomes^{7,19-21}. It has been argued that disturbance of *trans*-sensing mechanisms could lead to instability of chromosome pairing, a feature possibly also associated with some forms of cancer¹⁹.

The strong conservation of specific domains, including a putative zinc-finger, between *Psc*, *Su(z)2* and *Bmi-1* could indicate a molecular role for these domains in regulating gene expression by influencing higher-order chromatin structure. We are currently determining whether these domains are involved in specific protein-protein or protein-DNA interactions. □

Received 10 June; accepted 15 August 1991.

1. van Lohuizen, M. et al. *Cell* **65**, 737-752 (1991).
2. Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P. & Adams, J. M. *Cell* **65**, 753-763 (1991).
3. Tagawa, M. et al. *J. Biol. Chem.* **265**, 20021-20026 (1990).
4. Jürgens, G. *Nature* **316**, 153-155 (1985).
5. Adler, P. N., Martin, E. C., Charlton, J. & Jones, K. *Dev. Genet.* (in the press).
6. Brunk, B. P., Martin, E. C., Sharp, E. J. & Adler, P. N. *Nature* **353**, 351-353 (1991).
7. Wu, C. T., Jones, R. S., Lasko, P. F. & Gelbart, W. M. *Molec. gen. Genet.* **218**, 559-564 (1989).
8. Adler, P. N., Charlton, J. & Brunk, B. *Dev. Genet.* **10**, 249-260 (1989).
9. Brown, N. H. & Kafatos, F. C. *J. molec. Biol.* **203**, 425-437 (1988).
10. Brunk, B. P., Martin, E. C. & Adler, P. N. *Genetics* **128**, 119-132 (1991).
11. Brunk, B. P. & Adler, P. N. *Genetics* **124**, 145-156 (1990).
12. Paro, R. *Trends Genet.* **6**, 416-421 (1990).
13. Lewis, E. B. *Nature* **276**, 565-570 (1978).
14. Zink, B., Engström, Y., Gehring, W. J. & Paro, R. *EMBO J.* **10**, 153-162 (1991).
15. Zink, B. & Paro, R. *Nature* **337**, 468-471 (1989).
16. Paro, R. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **88**, 263-267 (1991).
17. Gaunt, S. J. & Singh, P. B. *Trends Genet.* **6**, 208-212 (1990).
18. Locke, J., Kotarski, M. A. & Tartof, K. D. *Genetics* **120**, 181-198 (1988).
19. Tartof, K. D. & Henikoff, S. *Cell* **65**, 201-203 (1991).
20. Wu, C. T. & Goldberg, M. L. *Trends Genet.* **5**, 189-194 (1989).
21. Bickel, S. & Pirotta, V. *EMBO J.* **9**, 2959-2967 (1990).
22. Higgins, D. G. & Sharp, P. M. *Gene* **73**, 237-244 (1988).

ACKNOWLEDGEMENTS. We thank R. Nusse and E. Wagenaar for libraries and RNA, C. Nüsslein-Volhard and members of her group for their hospitality, H. te Riele, J. Domen, C. Saris and P. Borst for comments on the manuscript, and P. N. Adler, C. T. Wu and R. Paro for discussion and sharing unpublished information.

Structural and serological similarity of MHC-linked LMP and proteasome (multicatalytic proteinase) complexes

Michael G. Brown, James Driscoll* & John J. Monaco

Department of Microbiology and Immunology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298-0678, USA

* Department of Cellular and Molecular Physiology, Harvard Medical School, Boston, Massachusetts 02115, USA

MAJOR histocompatibility complex (MHC) class I molecules associate with peptides derived from endogenously synthesized antigens. Cytotoxic T-lymphocytes can thus scan class I molecules and bound peptide on the surface of cells for foreign antigenic determinants. Recent evidence^{1,2} demonstrates that the products of *trans*-acting, non-class I genes in the class II region of the MHC are required in the class I antigen-processing pathway. There are genes (called *HAM1* and *HAM2* in the mouse) in this region that encode proteins postulated to be involved in the transport of peptide fragments into the endoplasmic reticulum for association with newly synthesized class I molecules²⁻⁵. But, the mechanism by which such peptide fragments are produced remains a mystery. At least two genes encoding subunits of the low-molecular mass polypeptide (LMP) complex are tightly linked to the *HAM1* and *HAM2* genes. We show that the LMP complex is closely related

to the proteasome (multicatalytic proteinase complex), an intracellular protein complex that has multiple proteolytic activities^{6,7}. We speculate that the LMP complex may have a role in MHC class I antigen processing, and therefore that the MHC contains a cluster of genes required for distinct functions in the antigen processing pathway.

The LMP antigens⁸ associate noncovalently *in vivo* to form a large protein complex of relative molecular mass 580,000 (M_r , 580 K)⁹. Two-dimensional PAGE analysis resolves the LMP complex into 16 nonidentical polypeptide subunits with a range of M_r (~20 K-30 K) and isoelectric points ($pI \approx 4-9$); ref. 8). LMP antigens are nonglycosylated interferon- γ inducible, intracellular polypeptides biochemically, serologically and genetically distinct from MHC class I, II and III antigens¹⁰. The LMP complex has broad tissue distribution and is found in both mouse and human⁸⁻¹⁰. At least two of the LMP complex polypeptides are genetically linked to the murine MHC class II region, and one of these (LMP-2) maps between the class II *Pb* and *Ob* genes¹⁰. More recently LMP-2 subunit complementary DNA has been cloned¹¹ and more precisely mapped to within 10 kilobases (kb) of the *HAM1* gene³.

The unusual biochemical characteristics of the LMP complex are very similar to the macromolecular protein complex, proteasome (M_r , 650 K), found in all eukaryotic tissues examined (for review, see ref. 12). The proteasome consists of 15-20 noncovalently associated subunits whose M_s range from 21 K to 35 K^{12,13} and whose isoelectric points range from pH 4 to 9 (ref. 14). The majority of proteasome is found in the cytoplasm, although some is in the nucleus¹³. This complex has proteolytic activity and degrades a number of protein substrates *in vitro*^{12,13}. Proteasome is an essential component of the 26S (M_r , 1,500K) ubiquitin and energy-dependent proteolytic complex that is responsible for the degradation of proteins that have been conjugated to ubiquitin^{15,16}, although proteasome cleaves both ubiquitinated and nonubiquitinated protein substrates^{12,17,18}. Thus, the proteasome may be a major pathway of nonlysosomal intracellular protein breakdown^{12,13,19}.

The overall similarity of the LMP complex to proteasomes of other species, and the fact that gel filtration fractions of murine cells that would be expected to contain the proteasome have been shown to contain only LMP subunits in the appropriate M_r range⁹, suggested that these two structures might be identical. To examine this issue directly, we compared anti-LMP and anti-proteasome immunoprecipitates from a murine macrophage cell extract (Fig. 1). A congenic BALB.B anti-BALB/c antiserum⁸ (anti-H-2^d) was used to immunoprecipitate LMP antigens, and a rabbit antiserum made against purified rat liver proteasomes (see Fig. 1 legend) was used to immunoprecipitate murine proteasomes (the proteasome is known to dissociate from the larger 26S ubiquitin-dependent proteolytic complex under the conditions used). The LMP precipitate contains the characteristic set of polypeptide subunits⁸⁻¹⁰. But there is one additional polypeptide not previously identified in LMP immunoprecipitates (Fig. 1a). This additional subunit may have been found because of increased resolution in M_r separation in our experiments, or because of limited proteolysis during purification. This subunit is also found in proteasome immunoprecipitates. The anti-proteasome precipitate yields a very similar pattern of subunits (Fig. 1b). Mixing of LMP and proteasome immunoprecipitates before electrophoresis (Fig. 1c) demonstrates that the 16 LMP subunits migrate with identical mobility to 16 of the polypeptides in the anti-proteasome precipitate. Interestingly, proteasome possesses the typical LMP H-2^d two-dimensional PAGE profile; two proteasome subunits migrate identically to the polymorphic, MHC-linked LMP-2 and LMP-7 polypeptides (Fig. 1), indicating that normal proteasomes contain both of these subunits. MHC class I molecules were not precipitated by anti-proteasome antiserum in contrast to anti-H-2^d precipitation (Fig. 1a, b). The anti-proteasome precipitate also contains four additional polypeptides not found

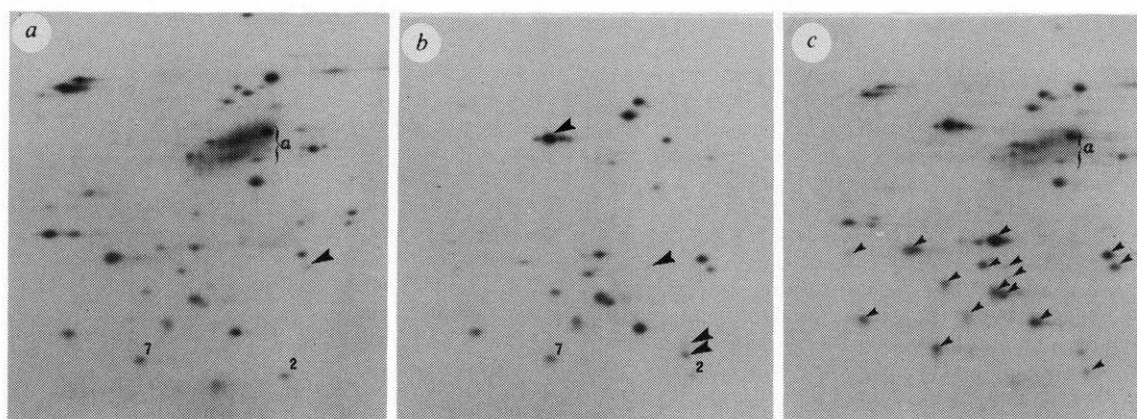


FIG. 1 Comparison of LMP and proteasome subunits by two-dimensional PAGE analysis. Immunoprecipitates from ^{35}S -methionine-labelled P388D1 extract with polyclonal mouse anti-H-2^d antiserum⁸ (a), polyclonal rabbit anti-proteasome (rodent) (b) and a mixture of the above immunoprecipitates (c). MHC class I heavy chains (α) are present in a and c but not in b. The polypeptide denoted by large arrow in a represents an LMP subunit not previously identified. LMP-2 and LMP-7, and the corresponding subunits in the anti-proteasome precipitate, have been denoted by numerals 2 and 7, respectively, in a and b. Polypeptides denoted by large arrows in b are found in proteasome immunoprecipitates but not in LMP immunoprecipitates. Small arrows in c indicate subunits of each complex that migrate with identical mobilities. All spots other than those marked are also found in control normal rabbit serum immunoprecipitates.

METHODS. Cell growth and labelling with ^{35}S -methionine, and immunoprecipitations were done as described previously⁸. Two-dimensional PAGE analysis was done as described previously⁸, except that 11% polyacrylamide gels were used for electrophoretic separation in the second dimension

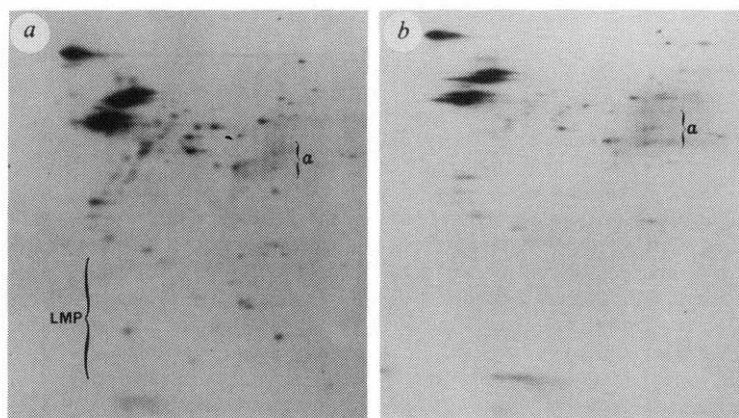
in LMP or control immunoprecipitates (Fig. 1b, large arrows). The remaining spots in both precipitates are also found in control (normal rabbit serum) immunoprecipitates (not shown).

These results suggest that the LMP complex shares most or all of its subunits with the proteasome. The additional polypeptides found in proteasome immunoprecipitates, but not in LMP immunoprecipitates might suggest that the two structures are not identical. These subunits may be selectively expressed only in the proteasome, such that the LMP complex may constitute only a portion of the larger proteasome. But it is also possible that the observed differences may be related to differences in the specificities of the antisera used. For example, both

instead of 10% polyacrylamide gels. Proteasome purified from rat liver as described previously¹⁷ was used to generate anti-proteasome antiserum. Enzyme purity was assessed by nondenaturing PAGE, coinciding protease activity against the substrate succinyl-Leu-Leu-Val-Tyr-MCA (MCA = 7-amido-4-methyl-coumarin), and SDS-PAGE. Purified enzyme emulsified in an equivalent volume of complete Freund's adjuvant was injected into a New Zealand White rabbit. Three weeks later a booster injection of purified enzyme emulsified in incomplete Freund's adjuvant was given. The rabbit was bled and serum obtained three weeks following booster injection. Anti-H-2^d antiserum was produced in BALB.B mice injected with BALB/c lymphoid cells as described⁸. Although this serum contains reactivity with class I and class II antigens, such antibodies are distinct from the antibodies which precipitate the LMP complex¹⁰. When antisera or monoclonal antibodies specific for either LMP, class I or class II antigens are used, no coprecipitation of class I and/or class II antigens with the LMP complex has ever been observed under the conditions used.

complexes may contain loosely associated subunits which dissociate from the complex under the conditions employed for immunoprecipitation. The anti-LMP antiserum probably recognizes a single polymorphic subunit¹⁰, whereas the rabbit anti-rat proteasome serum probably contains antibody to a large number of the polypeptides. Such polypeptides might then still be directly precipitated by the rabbit antiserum, but would be lost from anti-LMP precipitates. The same line of argument could be used to support the possibility that both LMP and proteasome contain unique as well as common subunits. In this case, anti-LMP antibody would precipitate only LMP, but both complexes would be precipitated by polyclonal anti-proteasome

FIG. 2 Anti-proteasome antiserum preclears WEHI-3 extract for LMP antigens. Immunoprecipitates with polyclonal anti-H-2^d antiserum from interferon (IFN)- γ -induced (20 U ml⁻¹) WEHI-3 extract (a), and IFN- γ induced (20 U ml⁻¹) WEHI-3 extract pre-cleared with anti-proteasome antiserum (b). LMP antigens are detected in a but not in b. MHC class I heavy chains (α) are present in both a and b. Immunoprecipitations and two-dimensional PAGE are described in the legend to Fig. 1. Anti-proteasome antiserum (90 μ l) was bound to protein A-bearing *Staphylococcus aureus* cells (Pansorbin, Calbiochem) in 10% suspension, divided into thirds and pelleted. Anti-H-2^d antiserum (120 μ l) was bound to *S. aureus* cells, divided in half and pelleted. One third of the anti-proteasome pellet was mixed with 100 μ l IFN- γ -induced WEHI-3 extract. The mixture was incubated at 4 °C for 25 min; these were standard conditions for subsequent incubations. The suspension was centrifuged and the pellet was collected, washed and eluted into non-equilibrium pH gradient electrophoresis (NEPHGE) sample buffer. The supernatant was mixed with another third of the anti-proteasome pellet and incubated under standard conditions. The pellet was collected and processed as above. The supernatant was then divided in half; one half was mixed with the remaining anti-proteasome pellet, the other half was mixed with one half of the anti-H-2^d pellet. Mixtures were incubated under standard conditions. Pellets were collected, washed and



eluted into NEPHGE sample buffer. The remaining half of the anti-H-2^d pellet was mixed with 50 μ l IFN- γ -induced WEHI-3 extract (not pre-cleared); the pellet was collected and processed as above. Each sample (25 μ l) was analysed by two-dimensional PAGE.

antibody by virtue of antibody directed against the common subunits. Alternatively, some or all of the four proteasome-specific polypeptides may represent polypeptides unrelated to the proteasome which coprecipitate with authentic proteasomes when this antiserum is used. In this regard, one of the four additional polypeptides, presumably having several isoforms different in pI (Fig. 1b), has a higher M_r than the characteristic subunit M_r range for either complex; however, precipitation of this polypeptide is inconsistent and therefore the significance of this finding remains undetermined.

To delineate further the relationship between LMP and proteasome, we tested the ability of anti-proteasome antiserum to pre-clear WEHI-3 extract for the LMP antigens (Fig. 2). Anti-proteasome antiserum effectively pre-cleared both for proteasome (not shown), and for LMP antigens, but not for MHC class I antigens (Fig. 2b), thus confirming the similarity of the proteasome and LMP complexes. Attempts to pre-clear proteasome from labelled extracts using our anti-LMP antiserum have been unsuccessful, although this antiserum does effectively pre-clear extract for the LMP antigens (not shown). This result suggests that the LMP complex differs serologically from the proteasome. But the biochemical basis of this serological difference is unclear, as proteasome precipitates from anti-LMP pre-cleared extract do not seem to differ from control proteasome precipitates. Possibly the polymorphic determinant recognized by anti-LMP may be concealed when the additional proteasome-specific subunits are bound to the complex. More detailed biochemical and serological analyses will be required to resolve these issues. But it is clear from these data that the LMP complex and proteasome are closely related structures.

Because of its broad tissue distribution and inducibility with interferon- γ , we previously suggested that the LMP complex might be involved in antigen processing¹⁰. The intracellular localization of the LMP complex is compatible with evidence demonstrating that class I antigen processing occurs non-lysosomally. Recently, cell lines that are defective in surface expression of MHC class I molecules owing to defects in antigen processing have been isolated, in which the genetic defects map to the class II region of the MHC (ref. 1). Two genes (*HAM1* and *HAM2*) were found in the murine MHC class II region whose amino-acid sequences indicate that they are members of a large family of transport proteins³. The products of the *HAM1* and *HAM2* genes therefore have been implicated in transporting peptide fragments from the cytoplasm into the endoplasmic reticulum where they may associate with newly synthesized MHC class I molecules²⁻⁵. Indeed, a cloned copy of the human homologue of *HAM1*, *Y3* (now called *PSF1*; ref. 2), corrects the class I defect in one mutant cell line²⁰. But the same cloned gene does not correct the defect in the deletion mutant .174, indicating that at least two (and perhaps more) closely linked genes in this region are required²⁰.

We have shown that the *LMP-2* gene is tightly linked to *HAM1*, and is homologous both to rat and to human proteasome subunits¹¹. Thus at least one LMP subunit maps to a genetic region important for antigen processing. We speculate that the role of the LMP complex may be to generate the peptides from cytoplasmic antigens that are transported into the endoplasmic reticulum by the *HAM* gene products. It is interesting that the protein found inconsistently in proteasome immunoprecipitates is similar in size and charge to that predicted for the *HAM1* gene product, as this might suggest physical interaction between these two structures.

The unusual structure of the LMP complex/proteasome may be relevant to its function. Recent analysis of the products of digestion of the insulin B chain by purified human proteasomes (macropain)²¹ is consistent with this hypothesis. First, discrete peptides, rather than free amino acids are produced. Second, although the structure contains two or more catalytic centres (with differential sensitivities to inhibitors), and the intact complex can cleave after basic, acidic or hydrophobic residues, a

single catalytic centre may have all of these reactivities, depending on the substrate. This observation suggests that the geometry of the complex may be more important than the chemical nature of the peptide bond in determining where cleavage occurs. Finally, kinetic analysis was done for one peptide that required two proteolytic events in order to be generated; this analysis indicated that both cleavages occurred without dissociation of the substrate from the complex. The resulting peptide was nine amino-acids long. MHC class I molecules bind nonapeptides 100-1,000-fold more efficiently than peptides of larger sizes²². If the results obtained for insulin B chain can be generalized to other substrates, selection for production of appropriately sized peptide may have pressured LMP molecules to coevolve with other MHC molecules involved in antigen processing and presentation, and may help to explain the genetic linkage between the genes encoding LMP antigens, the HAM transporter, and MHC class I molecules. Hence the LMP complex and the HAM transporter may be all that is required for antigen processing for class I-restricted immune responses. □

Received 31 May; accepted 9 August 1991.

1. Cerundolo, V. *et al. Nature* **345**, 449-452 (1990).
2. Spies, T. *et al. Nature* **348**, 744-747 (1990).
3. Monaco, J. J., Cho, S. & Attaya, M. *Science* **250**, 1723-1726 (1990).
4. Deverson, E. V. *et al. Nature* **348**, 738-741 (1990).
5. Trowsdale, J. *et al. Nature* **348**, 741-744 (1990).
6. Wilk, S. & Orłowski, M. *J. Neurochem.* **40**, 842-847 (1983).
7. Dahlmann, B., Kuehn, L., Rutschmann, M. & Reinauer, H. *Biochem. J.* **228**, 161-170 (1985).
8. Monaco, J. J. & McDavitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3001-3005 (1982).
9. Monaco, J. J. & McDavitt, H. O. *Nature* **309**, 797-799 (1984).
10. Monaco, J. J. & McDavitt, H. O. *Hum. Immun.* **15**, 416-426 (1986).
11. Martinez, C. K. & Monaco, J. J. *Nature* **353** (in the press).
12. Orłowski, M. *Biochemistry* **29**, 10289-10297 (1990).
13. Rivett, A. J. *Archs Biochem. Biophys.* **268**, 1-8 (1989).
14. Tanaka, K. *et al. J. molec. Biol.* **203**, 985-996 (1988).
15. Driscoll, J. & Goldberg, A. L. *J. biol. Chem.* **265**, 4789-4792 (1990).
16. Eytan, E., Ganoh, D., Armon, T. & Herskko, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7751-7755 (1989).
17. Driscoll, J. & Goldberg, A. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 787-791 (1989).
18. Matthews, W., Driscoll, J., Tanaka, K., Ichihara, A. & Goldberg, A. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2597-2601 (1989).
19. Arrigo, A. P., Tanaka, K., Goldberg, A. L. & Welch, W. J. *Nature* **331**, 192-194 (1988).
20. Spies, T. & DeMars, R. *Nature* **351**, 323-324 (1991).
21. Dick, L. R., Moomaw, C. R., DeMartino, G. N. & Slaughter, C. A. *Biochemistry* **30**, 2725-2734 (1991).
22. Schumacher, T. N. M. *et al. Nature* **350**, 703-706 (1991).

ACKNOWLEDGEMENTS. This work was supported by the NIGMS (J.J.M.).

A proteasome-related gene between the two ABC transporter loci in the class II region of the human MHC

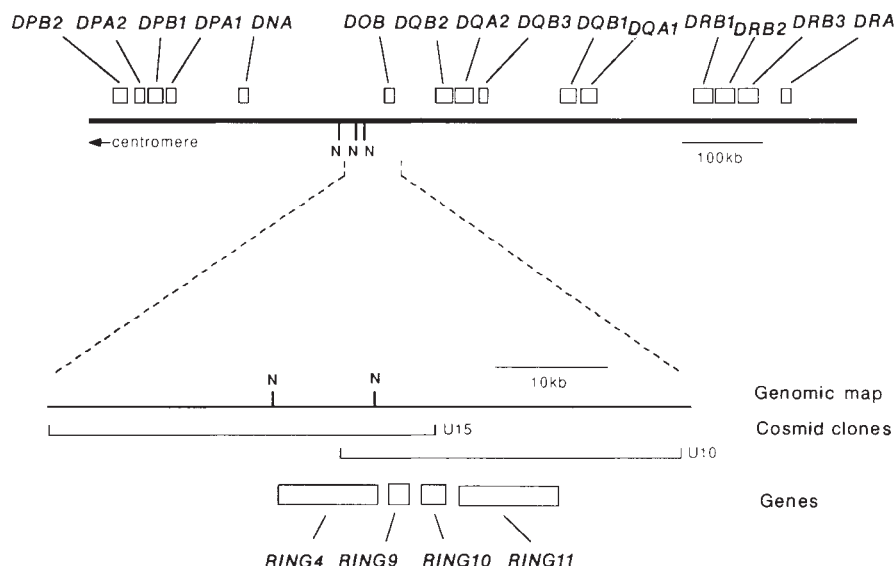
Richard Glynn, Stephen H. Powis, Stephan Beck, Adrian Kelly, Lesley-Anne Kerr* & John Trowsdale

Human Immunogenetics Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

* Molecular Immunogenetics, UMDS Guy's Campus, London SE1 9RT, UK

IT is now possible to paint a detailed picture of how cytoplasmic proteins are handled by the immune system¹⁻⁷. They are apparently degraded in the cytoplasm into peptides. These are then transported into the endoplasmic reticulum where they encounter class I major histocompatibility complex (MHC) molecules. Once loaded with peptide, the HLA molecules move through the Golgi apparatus to the cell membrane. Until recently, it had not been established how peptides without signal sequences cross the ER membrane. However, a number of papers have now described a pair of membrane transporter genes of the ABC (ATP-binding cassette) superfamily which are attractive candidates for this function⁸⁻¹². Both transporter genes, which may encode two halves of a heterodimer, are situated in the class II region of the MHC. There is evidence

FIG. 1 Location of the *RING4*, *9*, *10* and *11* genes in the class II region of the human MHC. The upper panel shows the main genes of the class II region on the short arm of human chromosome 6. The lower panel shows the positions of the *RING9* and *10* genes between the two *ABC* transporter genes, *RING4* and *11*. *N*, *NotI* restriction enzyme site. The map is based on a published map of the U15 and U10 cosmids¹⁰. The *RING4* and *11* genes probably correspond to *Y3* and *Y1*, respectively, but it is not known to which gene *Y2* corresponds¹⁰. The positions of *RING9* and *10* were determined by sequencing of genomic DNA. Mouse cosmid 5.9, which contains the murine transporter genes, also strongly hybridized to the *RING10* probe under non-stringent conditions (data not shown, ref. 11). Full details of the organization of the region at the genomic level will be published elsewhere (S.B. *et al.*, manuscript in preparation).



that other putative components of the processing machinery, the LMPs (low molecular mass polypeptides), are also encoded in the MHC¹³⁻¹⁵. Similarities between the properties of the LMPs and a large intracellular protease complex, called proteasome, have led to the suggestion that LMPs are involved in processing antigens¹⁶. We have now identified a human gene with sequence homology to proteasome components. Remarkably, this gene maps between the two putative peptide transporter genes.

By screening complementary DNA libraries for novel MHC genes, we obtained a number of expressed sequences in the class II region, designated the RING series. *RING4* (ref. 9) and *RING11* (S.H.P. *et al.*, manuscript in preparation) are two transporter genes mapping between *DNA* and *DOB*. These genes probably correspond to the *Y3* and *Y1* genes, respectively, described by Spies and coworkers¹⁰. cDNA libraries derived from a T-cell line, CEM, and a γ -interferon-stimulated monocyte line, U937, were screened with fragments of cosmids U10 and U15 (ref. 10). This revealed another two genes, *RING9* and *RING10*, located between *RING4* and *RING11* (Fig. 1).

The sequence of the longest *RING10* cDNA, isolated from the CEM library, is presented in Fig. 2. The sequence contains a long open reading frame which, from the first methionine, is predicted to encode a protein of 272 amino acids, relative molecular mass 30,000 (M_r , 30K) and pI 5.41. There is evidence to favour the second in-frame methionine at amino-acid position 65 as the translational start site: first, it matches the Kozak consensus more closely¹⁷ and second, it is only a few residues upstream of a region which shows sequence homology with a number of N-terminal sequences from rat and human proteasomes (see later and Fig. 3). This would result in a protein of 208 amino acids, M_r 23K and pI 7.78.

The derived amino-acid sequence of *RING10* was compared with sequences in the PIR 27, Swiss Prot 17 and OWL databases and a match found with a component of human proteasome. The proteasome (also known as macropain or multicatalytic proteinase complex) is a large intracellular protease. It consists of about 15 polypeptide components, depending on the species. These components range in M_r from 20 to 35K and are ordered

FIG. 2 Sequence of the *RING10* cDNA. The longest clone from a cDNA library made from the T-cell line CEM was sequenced using the chain-termination method of Sanger²⁸. The cDNA library was in a derivative of the CDM8 vector and was prepared as described by Seed²⁸. The open reading frame is shown encoding a protein of 272 amino acids (single-letter amino-acid code). $\wedge$, The potential initiating methionines. The predicted amino-acid sequence from the first AUG is shown although the context of this codon does not match the Kozak eukaryotic translation initiation sequence (CC^A/G^CCAUGG(G)) particularly well¹⁷. The second AUG is in slightly better context and may, in fact, represent the true N terminus of the protein (see text and Fig. 3). The consensus sequence for an active-site histidine from the

```

GGGCAGAAAGGGCAGCTCTTGTGGGTGACTACAGGTTAGGAGACCGTTGAACCTGGAGGGGCCCTAGGATGGACCCCGTGGAAGATTCAGAGACTGCG 100
CCCTCTCCCTGCGCCGCCCTTCCCTACACGCGGGGTATATCTGTGTGAGTGGCCAGGACCTGTTTCCAAGACTCTGCCCCCTCGCACTTCCGTC 200
CCTCCTGGTTTGTAAAGTGATGCTCATAGGAACCCCAACCCCGGTGACACTACTCCAGCTCTGGGCTGACTTCTAGTCTTCTGGTGAAGCTGCGGC 300
      M L I G T P T P R D T T P S S W L T S S L L V E A A P 27
TTTAGATGACACGACCTACCCACCCCTGTTTCCAGCGGATGCCCGGGCTGGAGCCACAGAAATCTTCCAGTCCCTGGGTGGGACGGGAGAAAGAAC 400
L D D T T L P T P V S S G C P G L E P T E F F Q S L G G D G E R N 60
GTTGAGTGGATGGCCCTGACCAACCCAGCTCGCCCTCAAGTCCAGCATGGAGTATTGACAGCTGGATCTCGGGCCTCAGCTGGGCTCTACA 500
V Q I E M A H G T T T L A F K F Q H G V I A A V D S R A S A G S Y I 94
TTAGTGCTTACGGGTGAACAAGGTGATTGAGATTAAACCTTACCTGCTTGGCACCATGTCTGGCTGTGCAGCAGACTGTCAGTACTGGGAGCGCTGCT 600
S A L R V N K V I E I N P Y L L G T M S G C A A D C Q Y W E R L L 127
GGCAAGGAATGCAGGCTGACTATCTGCGAAATGGAGAAGCTATTTTCAGTGTGCGGACGCTCCAAGCTGTGTCCAACATGATGTCCAGTACCGGGGC 700
A K E C R L Y Y L R N G E R I S V S A A S K L L S N M M C Q Y R G 160
ATGGCCTCTCTATGGGCAGTATGATCTGTGGCTGGGATAAGAAGGGTCTGGACTCTACTACGTGGATGAACATGGGACTCGGCTCTCAGGAATATGT 800
M G L S M G S M I C G W D K K G P G L Y Y V D E H G T R L S G N M F 194
TCTCCAGGGTAGTGGGAACACTTATGCTACGGGGTTCATGGACAGTGGCTACGGCCTAATCTTAGCCCTGAAGAGCCCTATGACCTTGGCCGAGGGC 900
S T G S G N T Y A Y G V M D S G Y R P N L S P E E A Y D L G R R A 227
TATTGCTTATGCCACTCAGAGACAGTATTCTGGAGGCGTGTGCAATATGTACCATGAAGGAAGATGGTGGGTGAAGTAGAAGTACAGATGTC 1000
I A Y A T H R D S Y S G G V V N M Y H M K E D G W V K V E S T D V 260
AGTGACCTGCTGCACCACTACCGGGAAGCCAATCAATAAGTGGTGGTGGCAGCTGGCAGGTCTCCTCTGGGAGGTCTGGCCGACTCAGGGACCTAA 1100
S D L L H Q Y R E A N Q * 272
GCCAGCTTAAGTCAAGGAGAAGAAGAGGCCCTAGCCTGAGCCAAAGAGAGATAC...

```

subtilisin family of serine proteases is underlined. The sequence is truncated at the 3' end and contains a further 124 nucleotides before the poly(A) tail.

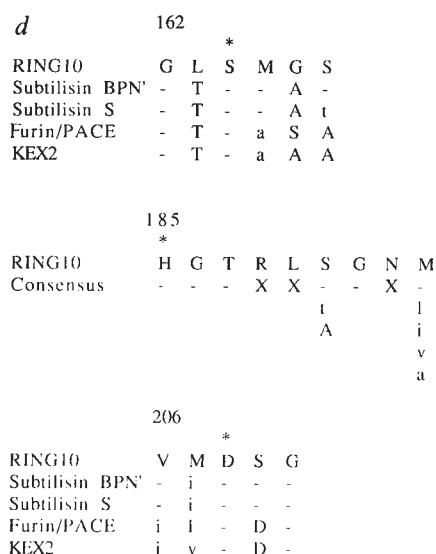
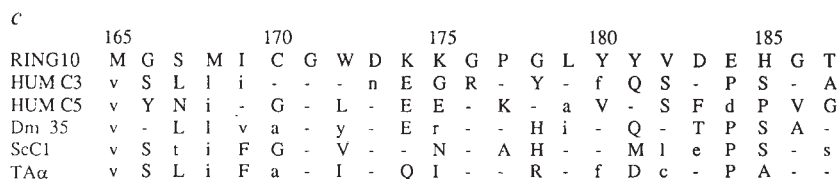
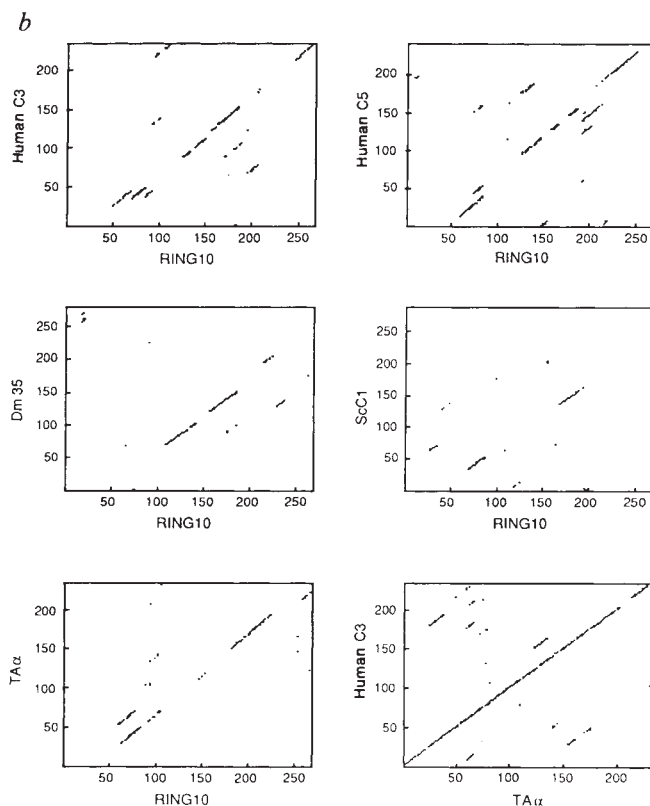
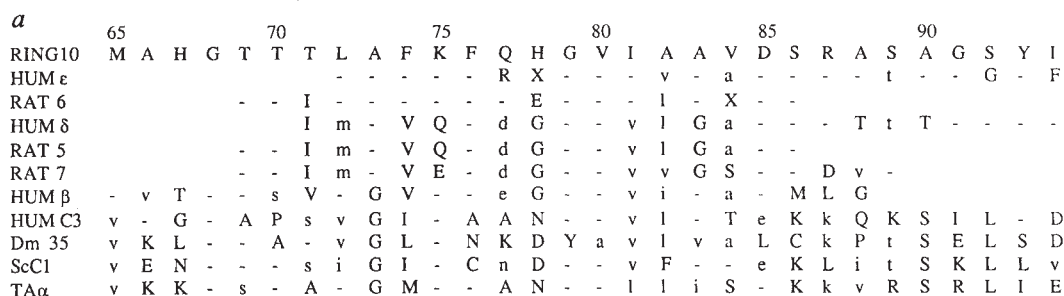


FIG. 3 Comparison of RING10 with proteasome components. *a*, Alignments with sequences derived from N-terminal amino-acid sequencing of human (HUM ϵ , HUM δ , HUM β ; ref. 19) and rat (RAT 5, 6, and 7; ref. 21) proteasome components and cDNA sequencing of human (HUM C3), *Drosophila melanogaster* (Dm35), *Saccharomyces cerevisiae* (ScC1) and *Thermoplasma acidophilum* (TA α) proteasome components^{22,23}. Dashes show positions at which there is identity with RING10, lower case indicates conservative changes³⁰. X, a residue for which identification was uncertain. Numbering refers to RING10. The RING10 and human ϵ chain sequences shared 15 out of 22 amino acids plus some conservative substitutions, indicating that the sequences are undoubtedly related. The alignment with cDNA-derived protein sequence is not as striking as that with the human ϵ or RAT 6 peptides. Note that, although the amino-acid sequences were obtained from N-terminal sequencing, they align at some distance from the first methionine of RING10. Although this may be due to proteolysis, it is consistent with translation initiating at the second in-frame AUG codon (see text and *b*). *b*, Diagon plots of cDNA-derived protein sequence from: Human C3, Human C5, Dm 35, ScC1 and TA α against RING10 and Human C3 against TA α . The comparisons were run using the Similarity Investigation Programme in the Staden package with a window of 31, a proportional score of 330, and the default score matrix. Note that extrapolation of the plots to the RING10 axis in the first five cases again suggests that the methionine at amino acid 65 is the initiating residue. The sixth diagan shows homology between Human C3 and TA α over the whole protein length for comparison. These two proteins have about 40% identical residues. The previously published full-length human chain sequences²³ show 10–35% identity with each other (data not shown). *c*, Alignment of RING10 with a stretch of Human C3, Human C5, Dm 35, ScC1 and TA α cDNA-derived protein sequences^{22,23}. Symbols and numbering are as in *a*. This alignment also shows that obtained with the N-terminal peptide sequence of RING10 with serine protease active sites. Symbols as in *a*, with the exception that X represents any amino acid. The serine, histidine and aspartate residues, which form the catalytic triad in RING10 are marked with an asterisk. The consensus, from the subtilisin family of serine proteases, screening the RING10 protein sequence using the program (ref. 26) and shows an exact match at amino acids 185–190. Sequences from known subtilisin-like serine proteases were used to mine this consensus. Alternative consensus residues are suggested. Suggested alignments, obtained by inspection, of residues surrounding the potential active site aspartate and serine, with those of known subtilisin-like serine proteases, stress that the involvement of any of these residues in the active site remains conjecture without relevant biochemical data. BPN' sequence is from *Bacillus amyloliquefaciens*, ssp. *var. amylosacchariticus*, furin is a human protease from *cerevisiae*^{27,31}.

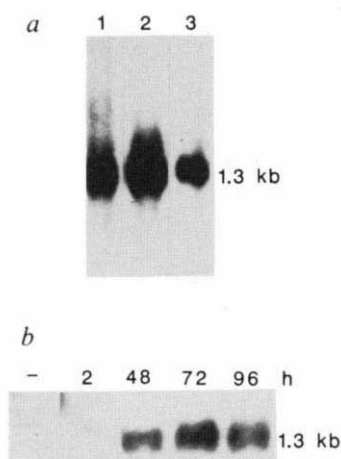


FIG. 4 Northern blots of *RING10* expression. *a*, Lanes 1, T-cell line (J6), 2, B-cell line (MANN) and 3, Human monocyte cell line (U937). *b*, Normal human keratinocytes were grown in 200 U recombinant human γ -interferon ml^{-1} for 2, 48, 72 and 96 h before RNA extraction (20 μg total cellular RNA was run per track). The first lane shows RNA from untreated cells. The cDNA clone shown in Fig. 2 was used as a probe. Autoradiography was for 24 h using Kodak XAR-5 film. Standard protocols were used³².

into a cylindrical structure of M_r 650K^{18–20}. A very good match was seen with a number of limited amino-acid sequences, obtained by N-terminal sequencing, from several human and rat proteasome components. Notable among these are human proteasome ϵ chain and rat chain 6 (Fig. 3*a* and refs 19, 21). Full-length comparisons with these chains await the cloning of their respective cDNAs. The alignments of *RING10* with derived protein sequences from other published cDNAs^{22,23} showed 10–20% identity. Low-stringency diagen plots between *RING10* and a selection of cDNA sequences from a variety of species are shown in Fig. 3*b*, and alignments of amino acids 65–94 and 165–187 are shown in Figs. 3*a* and *c*, respectively.

As the proteasome is thought to be a multiproteinase complex with broad specificity, its components should have active sites able to catalyse several proteolytic reactions²⁰. Although inhibition has been reported with both serine and cysteine protease inhibitors^{24,25}, homology to known consensus sequences for protease active sites has not been found in proteasome cDNAs²³. However, a protein motif search of the PROMOT database²⁶ revealed a good match between *RING10* and a consensus sequence derived from over twenty members of the subtilisin family of serine proteases. This suggests an active-site histidine at *RING10* amino acid 185 (Fig. 3*d*). The probability of such a match occurring randomly is close to 1 in 7,000.

Known serine protease active sites consist of three regions, each containing a residue involved in the catalytic mechanism. Typical consensus sites for the other two residues of the serine protease catalytic triad, aspartate and serine, were not identified by the motif search, but plausible alignments with known serine proteases could be found by inspection and are shown in Fig. 3*d*. The order of these putative catalytic residues in *RING10* (serine, histidine, aspartate) differs from that of both subtilisin-like proteases (aspartate, histidine, serine) and trypsin-like proteases (histidine, aspartate, serine)²⁷. Thus, if *RING10* is a serine protease, it is atypical and a definitive assignment of its function requires biochemical analysis.

The LMPs were originally defined by two-dimensional PAGE of immunoprecipitates obtained from a mouse macrophage line^{13–15}. They consist of 15 polypeptides with M_r s of 15–30K which noncovalently associate to form a complex of 580K. The complex resembles proteasome (see above and ref. 16), indicating that LMPs might be identical or related to proteasome components. Using recombinant strains, the genes for two of the LMPs have been localized to a region of the mouse MHC between *Pb* and *Ob*, the equivalents of *DNA* and *DOB* in the human (see Fig. 1). It is possible that other LMP genes are also located in the MHC but, due to a lack of detectable polymorphism, have not been mapped with this method. It has been proposed that LMPs might be involved in antigen processing,

degrading cytoplasmic proteins into antigenic peptides before their transport into the ER by the MHC-linked transporters^{15,16}. Consistent with this role is the increased expression of LMP proteins in the presence of γ -interferon¹⁵, a property shared with class I and II MHC antigens and the putative peptide transporters (ref. 9, S.H.P., unpublished observations). Expression of *RING10* is also induced by γ -interferon (Fig. 4).

Our data provide further evidence to link LMPs with proteasome. We have found a proteasome-like gene in the precise region of the MHC where LMP genes were predicted to reside. In addition, the *RING10* messenger RNA is of a size consistent with that of both proteasome and LMP components^{15,22,23} and, like the LMPs, is induced by γ -interferon. The exact relationship between proteasome, LMPs and *RING10* remains to be established. But the finding of such an intimate association between a proteasome gene and the two transporter loci implies a close relationship at the regulatory or functional level. □

Received 11 June; accepted 1 August 1991.

1. Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457–467 (1985).
2. Townsend, A. R. M. *et al.* *Cell* **44**, 959–968 (1986).
3. Bjorkman, P. J. *et al.* *Nature* **329**, 512–516 (1987).
4. Morrison, L. A., Lukacher, A. E., Braciale, V. L., Fan, D. & Braciale, T. J. *J. exp. Med.* **163**, 903–921 (1986).
5. Moore, M. W., Carbone, F. R. & Bevan, M. J. *Cell* **54**, 777–785 (1988).
6. Yewdell, J. W. & Bennink, J. R. *Cell* **62**, 203–296 (1990).
7. Weiss, S. & Bogen, B. *Cell* **64**, 767–776 (1991).
8. Deverson, E. V. *et al.* *Nature* **348**, 738–741 (1990).
9. Trowsdale, J. *et al.* *Nature* **348**, 741–744 (1990).
10. Spies, T. *et al.* *Nature* **348**, 744–747 (1990).
11. Monaco, J. J., Cho, S. & Attaya, M. *Science* **250**, 1723–1726 (1990).
12. Spies, T. & DeMars, R. *Nature* **351**, 323–324 (1991).
13. Monaco, J. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3001–3005 (1982).
14. Monaco, J. J. & McDevitt, H. O. *Nature* **309**, 797–799 (1984).
15. Monaco, J. J. & McDevitt, H. O. *Human Immun.* **15**, 416–426 (1986).
16. Parham, P. *Nature* **348**, 674–675 (1990).
17. Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
18. McGuire, M. J. & DeMartino, G. N. *Biochim. biophys. Acta* **873**, 279–289 (1986).
19. Lee, L. W. *et al.* *Biochim. biophys. Acta* **1037**, 178–185 (1990).
20. Rivett, A. J. *Archs Biochem. Biophys.* **268**, 1–8 (1989).
21. Lilley, K. S., Davison, M. D. & Rivett, A. J. *FEBS Lett.* **262**, 327–329 (1990).
22. Zwickl, P., Lottspeich, F., Dahlmann, B. & Baumeister, W. *FEBS Lett.* **278**, 217–221 (1991).
23. Tamura, T. *et al.* *Biochim. biophys. Acta* **1089**, 95–103 (1991).
24. Rivett, A. J. *J. biol. Chem.* **260**, 12600–12606 (1985).
25. Tanaka, K., Ichiyama, A., Waxman, L. & Goldberg, A. L. *J. biol. Chem.* **261**, 15197–15203 (1986).
26. Sternberg, M. J. E. *Computer Applications in the Biosciences* **7**, 257–260 (1991).
27. Barr, P. J. *Cell* **66**, 1–3 (1991).
28. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
29. Seed, B. *Nature* **329**, 840–842 (1987).
30. Argos, P. J. *Molec. Biol.* **193**, 385–396 (1987).
31. Fuller, R. S., Brake, A. J. & Thorner, J. *Science* **246**, 482–486 (1989).
32. Ausubel, F. M. *et al.* *Current Protocols in Molecular Biology* (Wiley Interscience, New York, 1987).

ACKNOWLEDGEMENTS. We thank D. Simmonds for generously providing cDNA libraries, G. Blanck for cosmid U10 and U15, R. Lovering for northern filters and A. Townsend, A. J. Rivett and J. Brannigan for useful discussions. S.H.P. is an MRC Clinician Scientist Fellow, L.-A.K. is supported by the Wellcome Foundation.

Myc rescue of a mutant CSF-1 receptor impaired in mitogenic signalling

Martine F. Roussel*†, John L. Cleveland‡, Sheila A. Shurtleff*§ & Charles J. Sherr*†§||

Departments of * Tumor Cell Biology and ‡ Biochemistry, and § Howard Hughes Medical Institute, St Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, Tennessee 38105, USA
† Department of Biochemistry, University of Tennessee College of Medicine, Memphis, Tennessee 38163, USA

THE colony-stimulating factor-1 receptor (CSF-1R) mediates its pleiotropic effects through the coupling of its ligand-activated tyrosine kinase to multiple intracellular effector proteins, whose combined actions determine the magnitude and specificity of the biological response. The interaction of cytoplasmic signalling molecules with CSF-1R is mediated in part by sequence motifs flanking sites of receptor tyrosine phosphorylation¹. Mutation of an autophosphorylation site at tyrosine 809 in the cytoplasmic domain of human CSF-1R does not significantly reduce its ligand-stimulated tyrosine kinase activity, binding to phosphatidylinositol 3-kinase, or induction of the immediate early response genes, *c-fos* and *junB* (ref. 2). Unlike cells bearing wild-type receptors, mouse NIH3T3 cells expressing mutant CSF-1R(Phe 809) were unable to grow in serum-free medium containing human recombinant CSF-1 and did not form colonies in semi-solid medium in its presence. CSF-1 induction of *c-myc* messenger RNA in these cells was impaired, but enforced expression of an exogenous *c-myc* gene restored their ability to proliferate in response to the growth factor. These studies demonstrate a receptor-mediated bifurcation of intracellular signal transduction pathways during the immediate early response and assign a central role for *c-myc* in CSF-1-induced mitogenesis.

Northern blotting revealed that CSF-1 stimulation of quiescent, serum-starved NIH3T3 fibroblasts expressing wild-type human CSF-1R induced the rapid accumulation of *c-fos*, *junB* and *c-myc* mRNAs, whereas cells expressing similar levels of CSF-1R(Phe 809) manifested a severely reduced *c-myc* response to the growth factor (Fig. 1). By contrast, *c-myc* mRNA was readily induced in both cell lines by serum, with maximal levels attained within 2–3 h after stimulation, followed by a persistent expression at lower levels^{3–5}. Based on the relative quantities of actin mRNA used as an internal control, induction of *c-fos* and *junB* mRNAs in the two cell lines was comparable. Expression of *max* mRNA, the heterodimeric partner of *myc* (ref. 6), was constitutive and was not regulated by CSF-1 or serum (data not shown).

Cells expressing wild-type human CSF-1R proliferated in serum-free medium containing human recombinant CSF-1 as the only exogenous growth factor and remained in exponential phase after reaching confluence (Fig. 2a)⁷, whereas those bearing the mutant receptor did not grow under these conditions (Fig. 2b). But, in medium containing 10% serum, cells expressing wild-type or mutant CSF-1R proliferated at the same rates, grew to similar densities, and were inhibited by contact. The density-dependent growth arrest of these cells in serum-containing medium correlates with *c-myc* downregulation⁸, but CSF-1 overrides this effect (Fig. 2a) and induces continuous *c-myc* expression (data not shown).

To determine whether their loss of responsiveness to CSF-1 might be due to an inability to express high levels of *c-myc* mRNA, cells bearing CSF-1R(Phe 809) were infected with a murine *c-myc*-containing retrovirus and seeded in semi-solid

medium in the presence or absence of purified human recombinant CSF-1. Uninfected cells expressing wild-type CSF-1R formed large colonies in CSF-1, but cells expressing CSF-1R(Phe 809) yielded only occasional clusters (<20 cells per cluster after 3 weeks) which did not continue to proliferate. Infection of cells expressing CSF-1R(Phe 809) with the *c-myc* retrovirus enabled them to generate CSF-1-dependent colonies (Fig. 3). CSF-1R-negative NIH3T3 cells infected with the *c-myc* retrovirus or CSF-1R(Phe 809)-positive cells infected with a control vector did not form colonies in agar whether CSF-1 was present or not. Similarly, *c-myc*-infected cells expressing wild-type or mutant forms of CSF-1R did not form colonies in the absence of human CSF-1. Fifteen CSF-1 responsive colonies derived from *c-myc*-infected cells expressing the mutant receptor were grown up and tested for *myc* RNA expression by northern blotting. All expressed large amounts of viral *c-myc* mRNAs, even when rendered quiescent by prolonged serum deprivation (Fig. 4). When five such subclones were retested for their ability to grow in agar, each yielded colonies only in the presence of CSF-1 at frequencies about 50% lower than those obtained with cells coexpressing both wild-type CSF-1R and the exogenous *c-myc* gene (Table 1).

Myc-infected cell lines expressing either wild-type or mutant CSF-1R grew at a faster rate and to a twofold higher saturation density in medium containing 10% serum, but they remained unable to proliferate in serum-free medium lacking CSF-1 and died faster than uninfected cells (Fig. 2). However, constitutive *c-myc* expression restored the CSF-1 dependent proliferative response of cells bearing CSF-1R(Phe 809) (Fig. 2b), although their growth rates remained slower than those of cells bearing the wild-type receptor (Fig. 2a).

Myc expression is positively regulated by many growth factors^{3–5,8} and can be induced by the constitutive or conditional

TABLE 1 Growth of CSF-1-stimulated cells in semi-solid medium

NIH3T3 cell line	Colony-forming efficiency in CSF-1 (% of 5×10^3 cells plated)
CSF-1R negative	
Uninfected	0.0
Vector alone	0.0
<i>Myc</i> infected	0.0
CSF-1R (wild-type)	
Uninfected	23
<i>Myc</i> clone 1	38
<i>Myc</i> clone 2	67
<i>Myc</i> clone 3	28
<i>Myc</i> clone 4	40
<i>Myc</i> clone 5	31
CSF-1R (Phe 809)	
Uninfected	0.0
Vector alone	0.0
<i>Myc</i> (mass infected)	9
<i>Myc</i> clone 1	25
<i>Myc</i> clone 2	18
<i>Myc</i> clone 3	19
<i>Myc</i> clone 4	23
<i>Myc</i> clone 5	16

Culture supernatants from NIH3T3 cells producing an infectious mouse *c-myc* retrovirus pseudotyped with amphotropic murine leukaemia virus²³ were used to infect receptor-negative or receptor-positive cells. Two weeks post-infection, cells were seeded in semi-solid medium in the presence or absence of CSF-1 (Fig. 3), and colonies were enumerated after 3 weeks. Representative colonies were expanded and retested in the same assay (Table). No colonies were obtained for any of the cell lines in the absence of CSF-1.

|| To whom correspondence should be addressed.

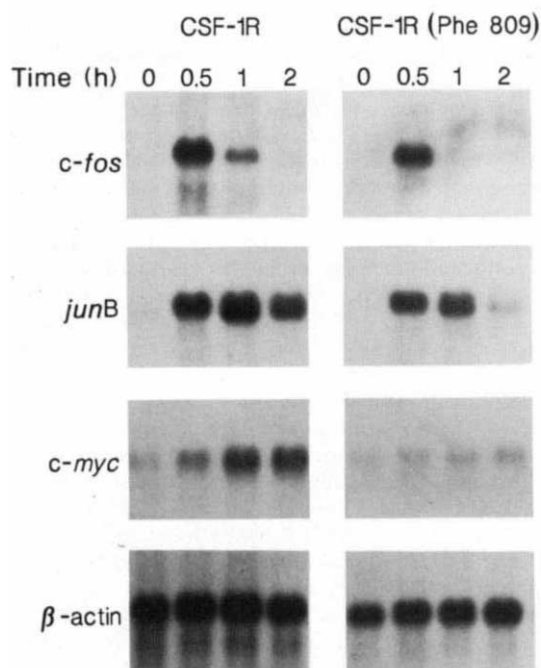


FIG. 1 Expression of immediate early response genes in CSF-1R-bearing NIH3T3 cells induced with CSF-1. NIH3T3 cells expressing wild-type (left) or mutant (right) CSF-1R were obtained by cotransfecting expression vectors containing human *c-fms* cDNAs together with pSV2neo, selecting neomycin-resistant cells in G418, and using fluorescence-activated flow cytometry (FACS) with a monoclonal antibody against human CSF-1R to sort receptor-positive populations^{22,24}. Cells were grown to confluence in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal calf serum (FCS) and then starved in DMEM containing 0.1% FCS for 48 h. Purified human recombinant CSF-1 was added to a final concentration of 1.8 nM for the indicated times. Total cellular RNA was extracted, separated electrophoretically on agarose gels containing formaldehyde, transferred to nitrocellulose, and hybridized with the indicated radiolabelled DNA probes in buffer containing 0.6 M NaCl and 50% (v/v) formamide at 42 °C for 18 h (ref. 23). Blots were washed in buffer containing 15 mM NaCl and 0.1% sodium dodecyl sulphate at 42 °C. Hybridizations were performed sequentially using randomly primed ³²P-labelled probes representing portions of *c-fos*, *junB*, *c-myc* and β -actin (ref. 9). Autoradiographic exposure times were 20 h for *junB*, *c-myc*, and β -actin and 3 days for *c-fos*. Using an LKB Ultrascan XL densitometer, the intensity of the *c-myc* bands in NIH3T3 cells expressing CSF-1R(Phe 809) represented less than 20% of that seen in cells expressing the wild-type receptor (three independent experiments).

FIG. 2 Growth of cell lines in the presence or absence of CSF-1. NIH3T3 cell lines expressing wild-type CSF-1R (a) or CSF-1R(Phe 809) (b) were plated at a density of 1×10^4 cells per 35-mm diameter culture dish in DMEM containing 10% FCS and starved after attachment for 48 h in medium containing 0.1% FCS. Replicate cultures were fed at 2-day intervals with (1) serum-free DMEM containing $400 \mu\text{g ml}^{-1}$ of bovine serum albumin and $10 \mu\text{g ml}^{-1}$ transferrin (serum free), (2) with serum-free medium containing 1.8 nM human recombinant CSF-1 (CSF-1), or (3) with DMEM containing 10% FCS (serum). Before feeding, parallel cultures were trypsinized, suspended in medium containing trypan blue vital dye, and the viable cells were counted. The results in a were obtained with uninfected cells expressing CSF-1R ($\square$, serum; $\bullet$, CSF-1; $\circ$, serum-free) or with a representative *myc*-infected subclone (clone 5, Table 1), ($\blacksquare$, serum + *myc*; $\blacktriangle$, CSF-1 + *myc*; $\triangle$, serum free + *myc*). The results in b were obtained with uninfected cells expressing CSF-1R(Phe 809) or with a representative *myc*-infected subclone (clone 4, Table 1) (symbols as in a).

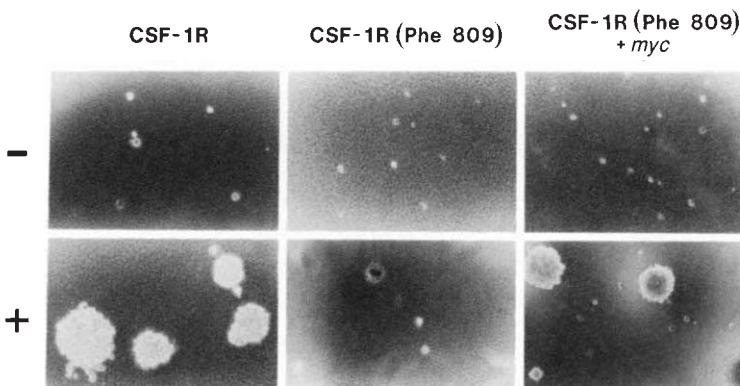
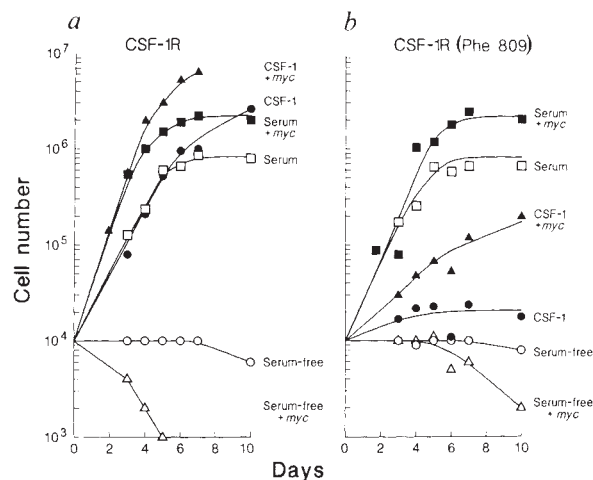


FIG. 3 Colony formation in semi-solid medium. NIH3T3 cells expressing wild-type or mutant CSF-1R were cloned in soft agar in Iscove's medium containing 15% FCS in the absence (-) or presence (+) of 1.8 nM purified human recombinant CSF-1 (ref. 24). Cells expressing the mutant receptor were infected with a *c-myc* retroviral vector pseudotyped with amphotropic murine leukaemia virus²³ and seeded as single cells in agar 2 weeks postinfection. 5×10^3 cells per ml were plated in each 35-mm diameter culture dish and colonies were photographed after 2 weeks. CSF-1-responsive colonies expressing CSF-1(Phe 809) and the viral *myc* gene were picked, expanded and retested (Table 1).

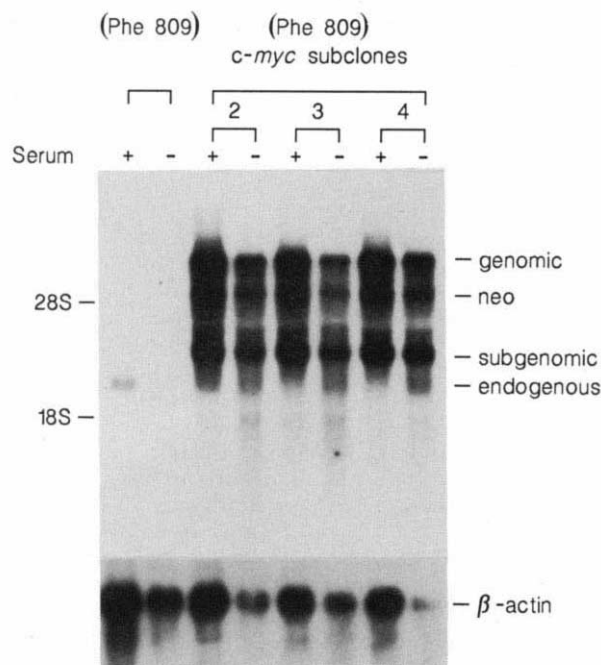


FIG. 4 Expression of *c-myc* retroviral construct in cells expressing CSF-1R(Phe 809). RNA was extracted from cultures expanded in 10% FCS (+) or starved of serum for 48 h (-), and 10- μ g samples were analysed for *c-myc* and β -actin mRNA by northern blot hybridization (Fig. 1). Exposure time for autoradiography was 20 h. The vector contains murine *c-myc* coding sequences (exons 2+3) inserted 3' to a splice acceptor sequence. Expression of the neomycin-resistance gene located 5' of *myc* is driven by an internal promoter, whereas spliced *myc* transcripts originate in the 5' viral long terminal repeat. The virus encodes three mRNA species corresponding to a full genome length transcript (genomic), a *neo-myc* mRNA (*neo*), and a spliced subgenomic *myc* species (subgenomic) which is the only mRNA to encode the *myc* protein²³. The position of the endogenous cellular *c-myc* mRNA is also indicated (endogenous). Migration of ribosomal RNA markers is indicated to the left.

expression of oncogenes with tyrosine kinase activity⁹. Under the latter circumstances, the action of *myc* in conferring mitogenicity has been difficult to evaluate, because many other genes are coinduced. More pertinent are experiments demonstrating that conditionally induced *myc* protein expression can bring rodent fibroblasts into cycle without affecting the expression of other immediate early genes, including *c-fos* (refs 5, 10). Conversely, *myc* antisense oligonucleotides can suppress the entry of cells into S phase^{11,12} as can the action of the retinoblastoma gene product, which negatively regulates *c-myc* expression¹³. The failure of ligand-stimulated CSF-1R(Phe 809) to induce *c-myc* in growth-arrested fibroblasts and the ability of exogenous *myc* expression to rescue their proliferative and colony-forming responses indicate that coupling between CSF-1R and *myc* is necessary for CSF-1-stimulated cell growth.

The signalling pathway leading from CSF-1R to *myc* represents one component of the immediate early response and can be distinguished from independent receptor-mediated events that determine the binding of phosphatidylinositol 3-kinase to

CSF-1R and the regulation of *c-fos* and *junB* expression². The latter responses may also be required for CSF-1-induced cell proliferation, but are clearly insufficient. What proximal effector(s) of the CSF-1R tyrosine kinase communicate with *c-myc*? Although the *raf-1* serine kinase neither binds to, nor is phosphorylated by, CSF-1R, its activity is indirectly induced by CSF-1 (ref. 14). Activation of *raf-1* can transactivate the *c-fos*, but not *c-myc* promoter^{15,16}, and both *v-raf* and *v-myc* synergize in transforming cells and conferring growth factor independence¹⁷⁻²⁰, suggesting that *c-raf* may function in a parallel pathway. Of the other known proximal effectors of growth factor receptor-induced responses, phospholipase C- γ 1 does not associate with CSF-1R (ref. 21) and p21^{ras} GTPase-activating protein (GAP) binds weakly and is, at best, phosphorylated at a very low stoichiometry following CSF-1 stimulation²². Therefore, an as-yet-uncharacterized effector which interacts with CSF-1R in the region surrounding tyrosine 809 may be required for *c-myc* induction and CSF-1-induced mitogenesis. □

Received 1 July; accepted 7 August 1991.

1. Cantley, L. C. *et al.* *Cell* **64**, 281-302 (1991).
2. Roussel, M. F., Shurtleff, S. A., Downing, J. R. & Sherr, C. J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6738-6742 (1990).
3. Kelly, K., Cochran, B., Stiles, C. D. & Leder, P. *Cell* **35**, 603-610 (1983).
4. Muller, R., Bravo, R., Burckhardt, J. & Curran, T. *Nature* **312**, 716-720 (1984).
5. Armelin, H. A. *et al.* *Nature* **310**, 655-660 (1984).
6. Blackwood, E. M. & Eisenman, R. N. *Science* **251**, 1211-1217 (1991).
7. Roussel, M. F. & Sherr, C. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7924-7927 (1989).
8. Dean, M. *et al.* *J. biol. Chem.* **261**, 9161-9166 (1986).
9. Cleveland, J. L., Dean, M., Rosenberg, N., Wang, J. Y. J. & Rapp, U. R. *Molec. cell. Biol.* **9**, 5685-5695 (1989).
10. Eilers, M., Schirm, S. & Bishop, J. M. *EMBO J.* **10**, 133-141 (1991).
11. Pieterpol, J. A., Holt, J. T., Stein, R. W. & Moses, H. L. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3758-3762 (1990).
12. Heikkilä, R. *et al.* *Nature* **328**, 445-449 (1987).

13. Pieterpol, J. A. *et al.* *Cell* **61**, 777-785 (1990).
14. Baccarini, M., Sabatini, D. M., App, H., Rapp, U. R. & Stanley, E. R. *EMBO J.* **9**, 3649-3657 (1990).
15. Jamal, S. & Ziff, E. B. *Nature* **344**, 463-466 (1990).
16. Siegfried, Z. & Ziff, E. B. *Molec. cell. Biol.* **10**, 6073-6078 (1990).
17. Rapp, U. R. *et al.* *J. Virol.* **55**, 23-33 (1985).
18. Blasi, E. *et al.* *Nature* **318**, 667-670 (1985).
19. Graf, T. *et al.* *Cell* **45**, 357-364 (1986).
20. Principato, M. *et al.* *Molec. cell. Biol.* **10**, 3562-3568 (1990).
21. Downing, J. R. *et al.* *EMBO J.* **8**, 3345-3350 (1989).
22. Reedijk, M., Liu, X. & Pawson, T. *Molec. cell. Biol.* **10**, 5601-5608 (1990).
23. Dean, M., Cleveland, J. L., Rapp, U. R. & Ihle, J. N. *Oncogene Res.* **1**, 279-296 (1987).
24. Roussel, M. F. *et al.* *Nature* **325**, 549-552 (1987).

ACKNOWLEDGEMENTS. We thank R. Eisenman (Fred Hutchinson Cancer Center) for a *max* probe, S. Clark (Genetics Institute) for human recombinant CSF-1, and C. Bockhold, B. Simmons and J. Watson for technical assistance.

The major chloroplast envelope polypeptide is the phosphate translocator and not the protein import receptor

Ulf-Ingo Flügge*, Andreas Weber*, Karsten Fischer*, Fritz Lottspeich†, Christoph Eckerskorn†, Karin Waagemann‡ & Jürgen Soll§

* Julius-von-Sachs-Institut für Biowissenschaften mit Botanischem Garten der Universität Würzburg, Mittlerer Dallenbergweg 64, 8700 Würzburg, Germany

† Max-Planck-Institut für Biochemie, Am Klopferspitz 18, 8033 Martinsried, Germany

‡ Botanisches Institut der Universität München, Menzinger Strasse 67, 8000 München 19, Germany

§ Botanisches Institut, Universität des Saarlandes, 6600 Saarbrücken, Germany

DURING photosynthetic CO₂ fixation, fixed carbon is exported from the chloroplasts in the form of triose phosphate by the chloroplast phosphate translocator, which is the principal polypeptide (E29) from spinach chloroplast envelopes¹. We have sequenced this nuclear-coded envelope membrane protein from both spinach and pea chloroplasts^{2,3}. An envelope membrane protein, E30, has been identified as a possible receptor for protein import into pea chloroplasts using an anti-idiotypic antibody approach⁴⁻⁶; antibodies raised against purified E30 inhibited binding and import of proteins into chloroplasts⁷. The amino-acid sequence of E30 deduced from its complementary DNA⁷ turned out to be highly homologous to that of E29, assigned by us as the spinach phosphate translocator², and was identical to the corresponding polypeptide from pea chloroplasts³. Differences in the binding properties to hydroxylapatite of E30 and the phosphate translocator suggested that E30 was not responsible for the chloroplast phosphate-transport activity but was the chloroplast import receptor⁷. Here we present evidence that argues against this and which identifies E30 as the chloroplast phosphate translocator.

Polypeptide profiles of envelope membranes from spinach and pea chloroplasts are shown in Fig. 1a. The main polypeptides of these membranes have different apparent relative molecular masses (*M_r*) of 29,000 (29 K; E29 (spinach); lane 1), and 30K (E30 (pea); lane 2). When the precursor proteins are synthesized from their cDNAs³ and processed *in vitro*, the mature forms have slightly different electrophoretic mobilities (Fig. 1a, lanes 6 and 7), which is consistent with the pea protein being larger by six amino-acid residues³. When intact chloroplasts from spinach or pea leaves are incubated with micromolar amounts of tritiated 4,4'-diisothiocyanostilbene-2,2'-disulphonic acid ([³H]DIDS), which strongly inhibits phosphate transport activity^{8,9}, the tritium label in the envelope membrane is associated with E29 and E30, respectively (Fig. 1a, lanes 4 and 5).

Other higher-molecular weight envelope polypeptides besides E29 are labelled when higher concentrations of [³H]DIDS are used and the resulting fluorogram is overexposed (Fig. 1a, lane 3).

To assess whether E29 is involved in phosphate transport we purified E29 and measured its activity in a reconstituted system¹⁰. We have used hydroxylapatite to purify solubilized phosphate-transport activity and found that it does not bind hydroxylapatite and so can be enriched by this procedure¹⁰. Subsequent chromatography of the unbound fraction containing this phosphate-transport activity on heparin-Sepharose CL-6B or chelating Sepharose 6B purified E29 to apparent homogeneity (Fig. 1b). In parallel, the specific reconstituted phosphate-transport activity increased (Table 1). Table 1 also shows that [³H]DIDS-labelled E29 can be purified 9-fold from envelope membranes, as determined from the [³H]DIDS/protein ratio. About 12% of the chloroplast envelope membrane protein therefore consists of E29. The reconstituted transport activity was enriched by only 5-fold, probably because of its functional instability in detergent¹⁰. Taken together these results demonstrate that E29 and [³H]DIDS-labelled E29, respectively, represent the phosphate translocator protein.

To confirm the identity of the [³H]DIDS-labelled E29 with the mature cDNA clone³, we sequenced a tryptic fragment of [³H]DIDS-labelled E29 from the amino terminus. The N-terminal amino-acid residues of a 26.4K fragment were ThrGlyPheLeuGluLysTyrProAla, corresponding to the sequence of the mature cDNA clone³ (amino-acid residues 11-19) and confirming the identity between them.

As shown in Fig. 1c and d, lanes 1-3, most of the functional translocator activity was not bound to hydroxylapatite at 4 °C and neither were E29 or [³H]DIDS-labelled E29, but the other solubilized envelope membrane proteins were largely absorbed and could only be eluted with high concentrations of phosphate. If the solubilized membrane proteins were exposed to hydroxylapatite for 30 min at room temperature as in the conditions used by Schnell *et al.*⁷, however, then most of the membrane proteins, including E29, remained bound to the column and could only be eluted with high concentrations of phosphate (Fig. 1c, lanes 4 and 5). The same holds for [³H]DIDS-labelled E29 (Fig. 1d, lanes 4 and 5). The absence of any immunoreactive E29 in this unbound fraction led to the conclusion⁷ that E29 was not the chloroplast phosphate translocator (expected in the unbound fraction) and could not therefore be responsible for chloroplast phosphate transport. But why no immunoreactive E29 was detected in the unbound fraction was because there was no E29 present in this fraction. The prolonged incubation at higher temperature probably favours denaturation of E29, causing it to be adsorbed by hydroxylapatite. Indeed, hydroxylapatite has been used to separate denatured and native membrane proteins^{11,12}. Presumably, the E29 (Fig. 1c), [³H]DIDS-labelled E29 (Fig. 1d), and E29 imported *in vitro* (Fig. 1e) which elute from hydroxylapatite only in high phosphate are partially denatured. This fraction also has a greatly reduced phosphate-transport activity.

TABLE 1 Isolation of phosphate-translocator activity and [³H]DIDS-labelled E29 from spinach chloroplast envelope membranes

Preparation	Protein (mg)	Reconstituted transport activity (nmol min ⁻¹)	Specific transport activity (nmol mg ⁻¹ min ⁻¹)	[³ H]DIDS-labelled E29 (d.p.m. 10 ⁻³ per mg)
Triton X-100-solubilized chloroplast envelope membranes	6.06	32.7	5.4	4.6
Chromatography on hydroxylapatite	0.50	8.4	16.8	30.1
Chromatography on heparin-Sepharose CL-6B or chelating Sepharose 6B	0.10	2.2	22.0	42.0

For experimental details see legend to Fig. 1b.

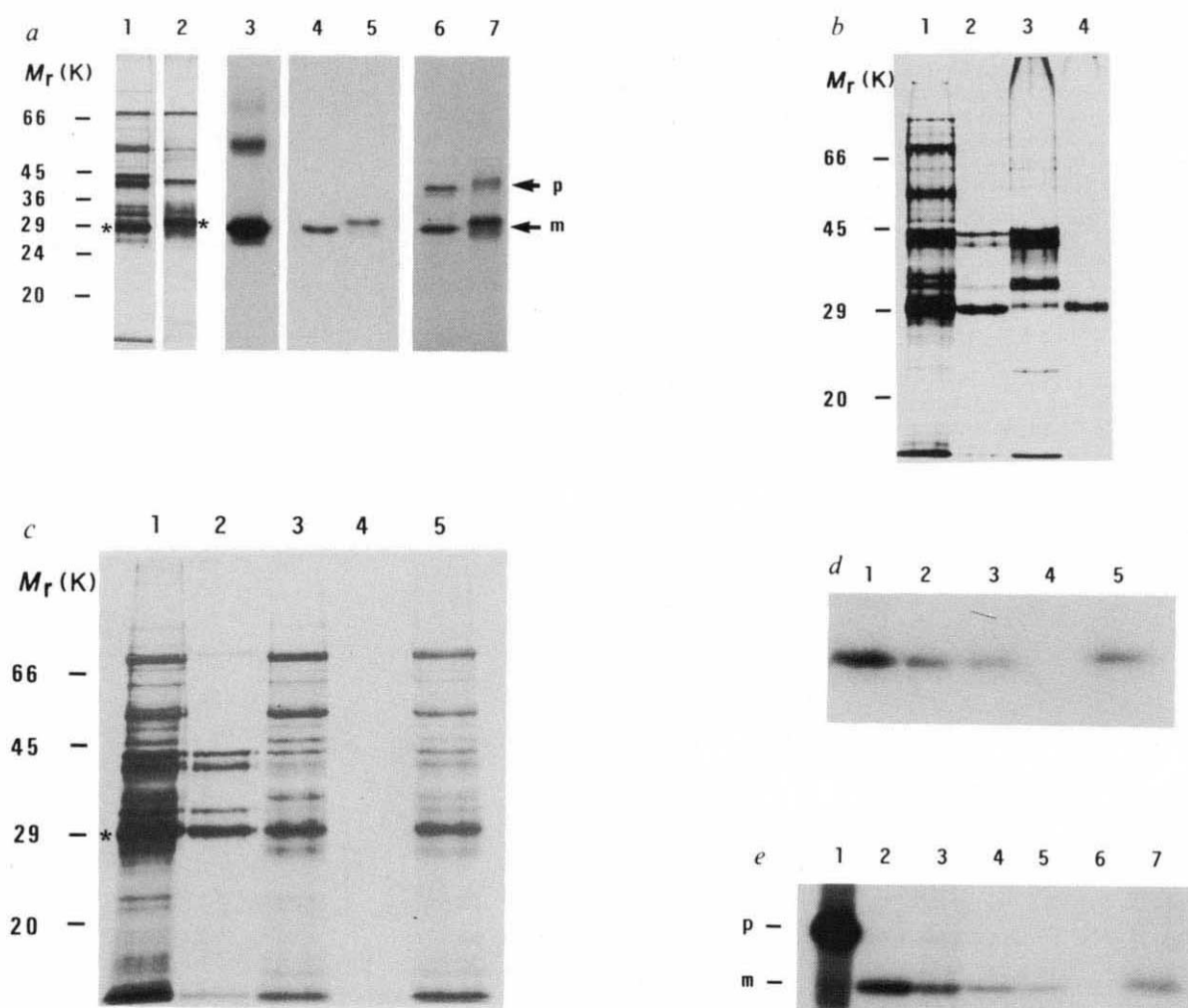
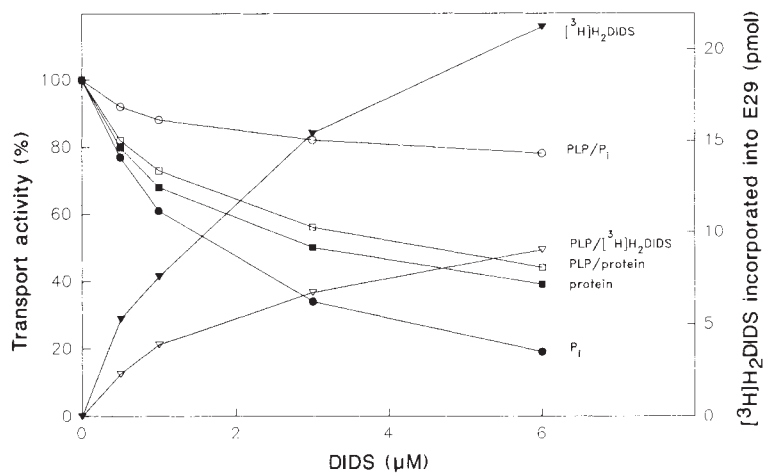


FIG. 1 The envelope membrane polypeptide E29 can be purified to homogeneity and does not bind to hydroxylapatite. *a*, SDS-PAGE of envelope membrane proteins from spinach (lanes 1, 3, 4) and pea (lanes 2, 5) chloroplasts, respectively. Lanes 1 and 2, silver-stained envelope membrane proteins E29 and E30, respectively, are marked by an asterisk; lanes 3–5, fluorographic analysis of [³H]DIDS-labelled envelope proteins: chloroplasts were incubated with 10 μ M [³H]DIDS (lane 3) or 2 μ M [³H]DIDS (lanes 4 and 5). Lanes 6 and 7, cDNA clones encoding E29 (spinach, lane 6) and E30 (pea, lane 7) were transcribed *in vitro*³ and translated in the presence of [³⁵S]methionine, processed *in vitro*³ and analysed by SDS-PAGE and fluorography; p and m represent the precursor and the mature forms of E29 and E30, respectively. *b*, Purification of E29 by chromatography on hydroxylapatite followed by chromatography on heparin-Sepharose CL-6B. Triton X-100-solubilized envelope membranes (lane 1) were fractionated on hydroxylapatite and the unbound fraction (lane 2) chromatographed on heparin-Sepharose CL-6B (lane 3, pass-through fraction; lane 4, fraction eluted with 0.6 M NaCl). *c* and *d*, Chromatography of solubilized envelope membranes (*c*, SDS-PAGE and silver staining; the position of E29 is marked by an asterisk) and solubilized [³H]DIDS-labelled envelope membranes (*d*, SDS-PAGE and fluorography) on hydroxylapatite at 4 °C (lanes 2 and 3) and 22 °C (lanes 4 and 5). Lanes 2 and 4 represent unbound fractions, lanes 3 and 5 fractions eluted from hydroxylapatite with 0.8 M sodium phosphate/0.1% SDS. *e*, Chromatography on hydroxylapatite of solubilized envelope membranes containing ³⁵S-labelled E29 imported *in vitro* which had been synthesized from the corresponding cDNA clone². Lane 1, precursor form of E29; lane 2, total envelope membrane proteins; lanes 3–5, fractions unbound to hydroxylapatite; lane 6, wash; lane 7, fraction eluted from hydroxylapatite with 0.8 M sodium phosphate/0.1% SDS. Fractions were analysed by SDS-PAGE and fluorography. In *a*–*c*, relative molecular mass calibration is shown on the left (in thousands).

METHODS. *a*, Intact chloroplasts from spinach and pea leaves were labelled by incubation with 2 μ M or 10 μ M [³H]DIDS⁹ and analysed by SDS-PAGE²⁶ with silver staining and fluorography²⁷. The cDNA clones encoding E29 and E30 were transcribed and translated *in vitro* as described³, processed *in vitro*³ and then subjected to SDS-PAGE and fluorography. *b*, Hydroxylapatite chromatography of purified envelope membranes was essentially as described¹⁰: envelope membranes (5 mg ml⁻¹) in 10 mM Tricine buffer, pH 7.4 were solubilized in Triton X-100 (final concentration, 3% v/v) for 2 min at 4 °C and added to hydroxylapatite (1.2 g per ml solubilized protein) that had been equilibrated with 10 mM Tricine buffer, pH 7.4, containing 0.2% Triton X-100 (buffer A). After 2 min on ice, the mixture was centrifuged (10,000g for 1 min) and the unbound fraction applied to a heparin-Sepharose CL-6B column (1.5 ml bed volume per ml sample; equilibrated with buffer A) which was eluted with a 0–1.0 M NaCl gradient. Aliquots of each fraction were assayed for protein²⁸, reconstitution of phosphate-transport activity¹⁰, and analysed by SDS-PAGE. *c* and *d*, Envelope membranes (*c*) or [³H]DIDS-labelled envelope membranes (*d*) were solubilized in Triton X-100 and incubated with hydroxylapatite for 5 min at 4 °C or for 30 min at 22 °C, respectively. The unbound fraction was collected by centrifugation and the hydroxylapatite washed three times with buffer A. The hydroxylapatite-bound proteins were eluted by washing the gel material with 0.8 M sodium phosphate, pH 7.2, 0.1% SDS and were then analysed by SDS-PAGE and silver staining (*c*) or by SDS-PAGE/fluorography (*d*). *e*, The ³⁵S-labelled precursor of E29 was synthesized by *in vitro* transcription/translation and imported into intact chloroplasts². The protein import buffer contained 0.05 M HEPES-KOH, pH 8.0, 0.3 M sorbitol, 10 mM methionine, 25 mM potassium gluconate, 0.2% bovine serum albumin, 50 mM lysine, 2 mM MgSO₄, 2 mM ATP. After 15 min at 22 °C, chloroplasts were treated with thermolysin (30 μ g ml⁻¹ for 15 min at 0 °C). Envelope membranes were isolated from the import assay², solubilized in Triton X-100 and chromatographed on hydroxylapatite.

FIG. 2 [³H]DIDS-labelling of E29 is linked to the inhibition of phosphate transport activity but not to protein import activity. *a*, Effect of pretreatment of the chloroplasts with pyridoxal-5-phosphate (PLP) on the inhibition of the transport of phosphate and on the inhibition of import of the ribulose-1,5-biphosphate carboxylase small subunit precursor (pSSU) by DIDS and, in addition, on the incorporation of [³H]DIDS into E29. Phosphate transport (circles) and import of pSSU (squares) into chloroplasts that had been pretreated without (filled symbols) and with PLP (open symbols); triangles represent incorporation of [³H]DIDS into E29 of chloroplast envelope membranes: ▼, control; ▽, effect of preincubation of the chloroplasts with PLP before the addition of [³H]DIDS. METHODS. *a*, Intact chloroplasts (0.17 mg chlorophyll ml⁻¹) were incubated in 0.33 M sorbitol, 10 mM tricine buffer, pH 7.4, with or without 0.08 mM PLP for 30 min at 4 °C and then for a further 30 min with increasing concentrations of DIDS or [³H]DIDS, respectively. They were then washed twice in 0.33 M sorbitol, 10 mM Tricine, pH 7.4, resuspended in protein import buffer (see legend to Fig. 1) and the phosphate transport activity²⁹, protein import activity and incorporation of [³H]DIDS into E29 (ref. 9), respectively, were measured after 30 min. For quantification of [³H]DIDS incorporated into E29 and of protein import activity, both E29 and the processed form of pSSU were cut out of and eluted from



the unstained SDS-polyacrylamide gel by incubation with 5% SDS for 12 h at 37 °C (ref. 9). Aliquots of centrifugal supernatants were used for liquid scintillation counting.

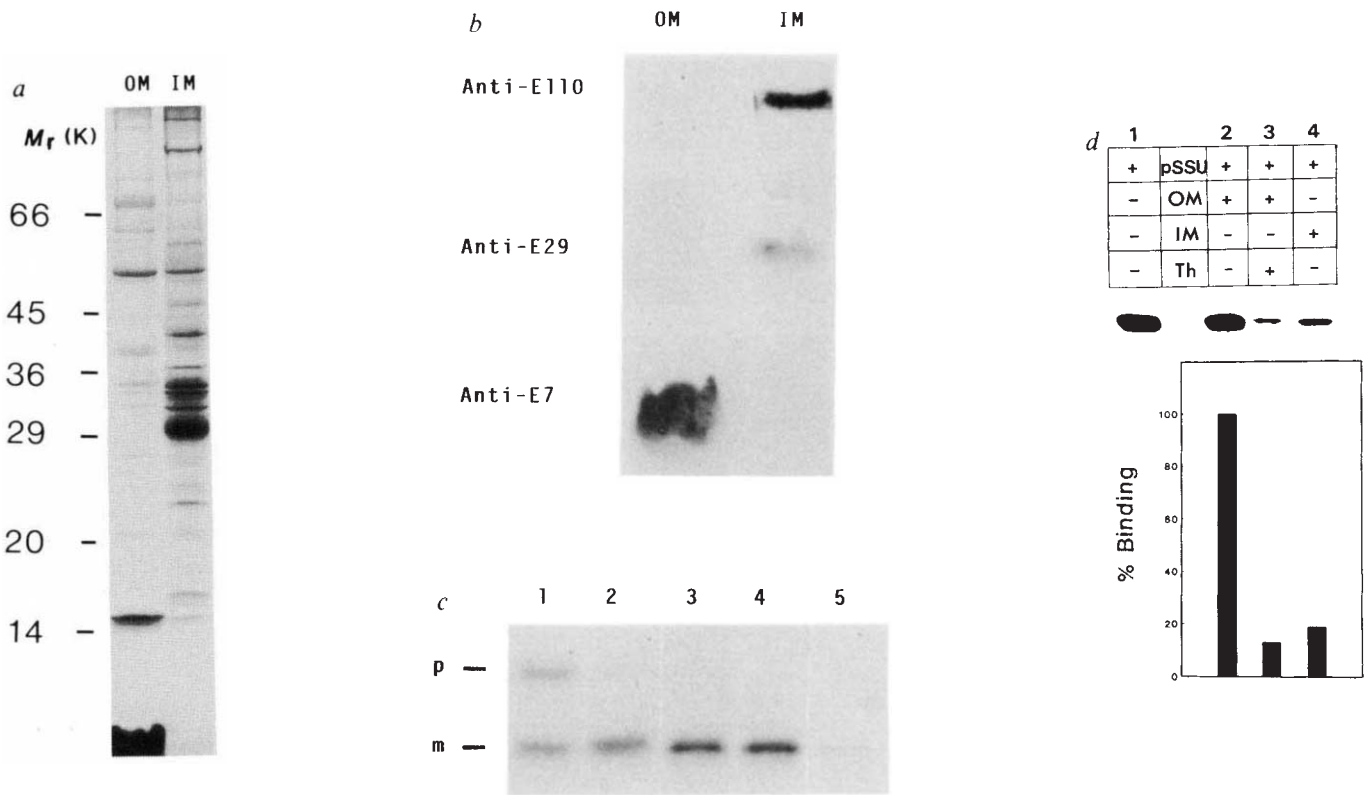


FIG. 3 E29 is a protease-resistant component of the inner envelope membrane but the protease-sensitive precursor binding activity is located in the outer envelope membrane. *a*, SDS-PAGE of purified outer (OM) and inner envelope membranes (IM) from pea chloroplasts. Relative molecular mass calibration shown on the left. Purified outer and inner envelope membranes were obtained by hypertonic lysis of chloroplasts followed by separation of the membrane vesicles on sucrose density gradients^{19,30}. *b*, Western-blot analysis of purified inner and outer envelope membranes probed with antibodies against E110 (an inner envelope membrane protein¹⁷), E29, and E7 (an outer envelope membrane protein¹⁸). *c*, Effect of thermolysin-treatment of chloroplasts (30 μg ml⁻¹ for 15 min at 4 °C) on *in vitro* imported

³⁵S-labelled E29 (lanes 1 and 2) and on [³H]DIDS-labelled E29 (lanes 3 and 4). Envelope membranes were isolated from the assays and analysed by SDS-PAGE and fluorography. Lane 5, before SDS-PAGE, envelope membranes containing imported E29 were treated with Triton X-100 (2%, v/v) in the presence of thermolysin (30 μg ml⁻¹ for 15 min at 4 °C). Precursor and mature forms of E29 are indicated by p and m respectively. *d*, Histogram showing that outer envelope membranes specifically bind precursor proteins in a protease-sensitive way. Th, thermolysin pretreatment. Assay numbers are indicated at the top; + and -, presence or absence of components in the assay. Binding of pSSU synthesized *in vitro* to outer and inner envelope membranes was as described¹⁹.

Figure 1e shows that E29 synthesized from its cDNA and imported *in vitro*, shares the properties of chloroplast envelope E29 [³H]DIDS-labelled E29, and the phosphate-transport activity: for the most part it is not bound to hydroxylapatite at 4 °C after solubilization in Triton X-100. These findings indicate that the cDNA clone reported in refs 2, 3 and 7 encodes the chloroplast phosphate translocator.

Figure 2 shows that DIDS inhibits the transport of both phosphate and protein, for example, the import of ribulose-1,5-bisphosphate carboxylase small subunit precursor. Increasing amounts of [³H]DIDS are incorporated into E29 as a function of [³H]DIDS concentration. Chloroplasts were alternatively pretreated with 0.08 mM pyridoxal-5'-phosphate (PLP), a potent inhibitor of phosphate transport¹³ which affects protein import activity only at higher concentrations (not shown), and incubated with increasing amounts of DIDS or [³H]DIDS. Subsequent treatment with lysine to hydrolyse the Schiff-base formed between PLP and the target protein, and to restore any PLP-inhibited activity, revealed that inhibition of protein import by DIDS was almost unchanged, whereas the inhibition of phosphate transport as well as the incorporation of [³H]DIDS into E29 were both drastically reduced. These data indicate that the incorporation of [³H]DIDS into E29 is linked to the inhibition of the transport of phosphate by DIDS and not to the protein import process. Thus, E29 is not involved in the translocation of proteins but rather of phosphate. One of the higher-molecular weight polypeptides that was labelled at higher [³H]DIDS concentrations and was revealed by overexposing the fluorogram (Fig. 1a) might possibly be associated with protein import function.

We next addressed the localization of E29 in the envelope membranes. Metabolite translocator proteins like the phosphate translocator reside in the inner envelope membrane, whereas receptor proteins are thought to be in the outer envelope. This agrees with the observation that pretreatment of chloroplasts with the protease thermolysin abolishes binding of precursor proteins¹⁴⁻¹⁶. Schnell *et al.*⁷, however, did not determine whether *in vitro*-synthesized E29 (their putative import receptor) was imported into the outer membrane or test whether it was protease-sensitive as would be expected for a receptor protein, nor did they establish whether antibody reactivity was restricted to the outer envelope. Figure 3a and b shows that E29 (and the 110K protein, an inner envelope membrane protein¹⁷) is exclusively localized to the inner envelope membrane. On the other hand, E7, an outer envelope membrane protein^{17,18}, could be

detected only in the outer envelope fraction. In line with this is the observation that only the surface-bound precursor form of *in vitro*-imported E29 is protease-sensitive, whereas the mature form of *in vitro*-imported E29 and [³H]DIDS-labelled E29 are both protease-resistant (Fig. 3c). In its solubilized form, however, E29 is completely digested by thermolysin (Fig. 3c). The localization of E29 to the inner envelope membrane was also demonstrated by western-blot analysis of envelope membranes from chloroplasts pretreated with thermolysin and from untreated controls (not shown).

Our localization of E29 could be attributed to the fractionation procedure of the inner and outer envelope membranes causing E29 to be pushed from putative contact sites between the inner and outer membranes into the inner envelope. We therefore investigated which membrane has the ability to bind precursor proteins. Figure 3d shows that in a functional envelope import system¹⁹, only the outer envelope membrane (not containing E29) can bind precursor proteins in a protease-dependent way, which further indicates that E29 is not involved in protein receptor function. The observation that antibodies directed against E29 react with components associated with precursor binding activity⁷ could be explained if the import receptor and the chloroplast phosphate translocator were to share common epitopes recognized by the antibody. This idea is supported by the observation that the mitochondrial 38K protein MOM38, which is part of the mitochondrial translocation apparatus, contains a sequence of ~70 amino acids with significant homology to an abundant 32K component of the inner mitochondrial membrane, the phosphate/OH⁻ carrier²⁰⁻²². This protein has been assigned as the mitochondrial import receptor using an anti-idiotypic approach^{23,24}. It is feasible that in chloroplasts E29 as an integral constituent of the inner envelope membrane could also have regions homologous to components of the translocation apparatus such as the chloroplast import receptor. It is possible that some antibodies raised against abundant membrane proteins (for example, E29 or the mitochondrial phosphate/OH⁻ carrier) might also recognize less abundant import receptors and so affect protein translocation into organelles. It is unlikely that an import receptor would be a major constituent of the chloroplast envelope membrane like E29: assuming 3,000-5,000 precursor binding sites per chloroplast²⁵, the import receptor should correspond to about 0.02-0.04% of the envelope membrane protein, which is far less than the amount of E29 in the membrane (~12%). □

Received 5 June; accepted 5 August 1991.

- Flügge, U.-I. & Heldt, H. W. in *Function and Molecular Aspects of Biomembrane Transport* (eds Quagliariello E., Palmieri F. & Klingenberg E. M.) 373-382 (Elsevier/North Holland, Amsterdam 1979).
- Flügge, U.-I. *et al.* *EMBO J.* **8**, 39-46 (1989).
- Willey, D. L., Fischer, K., Wachter, E., Link, T. A. & Flügge, U. I. *Planta* **183**, 451-461 (1991).
- Pain, D., Kanwar, Y. S. & Blobel, G. *Nature* **331**, 232-237 (1988).
- Joyard, J. & Douce, R. *Nature* **333**, 306-307 (1988).
- Pain, D. & Blobel, G. *Nature* **333**, 307 (1988).
- Schnell, D. J., Blobel, G. & Pain, D. *J. Cell Biol.* **111**, 1825-1838 (1990).
- Rumpha, M. E., Edwards, G. E., Yousif, A. E. & Keegstra, K. *Plant Physiol.* **86**, 1193-1198 (1988).
- Gross, A., Brückner, G., Heldt, H. W. & Flügge, U. I. *Planta* **180**, 262-271 (1990).
- Flügge, U. I. & Heldt, H. W. *Biochim. biophys. Acta* **638**, 296-304 (1981).
- Klingenberg, M. in *Membranes and Transport* (ed Martinosi, A. N.) Vol. 1, 203-209 (1982).
- Lin, C. & Klingenberg, M. *Biochemistry* **21**, 2950-2956 (1982).
- Flügge, U.-I. & Heldt, H. W. *FEBS Lett.* **82**, 29-33 (1977).
- Cline, K., Werner-Washburne, M., Andrews, J. & Keegstra, K. *Plant Physiol.* **75**, 675-678 (1984).
- Cline, K., Werner-Washburne, M., Lubben, T. H. & Keegstra, K. *J. Biol. Chem.* **260**, 3691-3696 (1985).

- Hinz, G. & Flügge, U. I. *Eur. J. Biochem.* **175**, 649-659 (1988).
- Block, M. A., Dorne, A. J., Joyard, J. & Douce, R. *J. Biol. Chem.* **258**, 13273-13280 (1983).
- Salomon, M., Fischer, K., Flügge, U. I. & Soll, J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5778-5782 (1990).
- Waegemann, K. & Soll, J. *The Plant J.* (in the press).
- Kiebler, M. *et al.* *Nature* **348**, 610-616 (1990).
- Zara, V. *et al.* *Eur. J. Biochem.* **198**, 405-410 (1991).
- Phelps, A., Schobert, C. & Wohlrab, H. *Biochemistry* **30**, 248-252 (1991).
- Pain, D., Murakami, H. & Blobel, G. *Nature* **347**, 444-449 (1990).
- Murakami, H., Blobel, G. & Pain, D. *Nature* **347**, 488-491 (1990).
- Friedman, A. L. & Keegstra, K. *Plant Physiol.* **89**, 993-999 (1989).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 84-88 (1974).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. Biol. Chem.* **193**, 265-275 (1951).
- Heldt, H. W. *Meth. Enzym.* **69**, 604-613 (1980).
- Keegstra, K. & Yousif, A. E. *Meth. Enzym.* **118**, 316-325 (1986).

ACKNOWLEDGEMENTS. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. K.F. was supported by the Studienstiftung des deutschen Volkes.

Antibiotic inhibition of group I ribozyme function

Uwe von Ahsen, Julian Davies* & Renée Schroeder

Institut für Mikrobiologie und Genetik der Universität Wien,
Althanstrasse 14, 1090 Vienna, Austria

*Unité de Génie Microbiologique, Institut Pasteur, 28 Rue du Dr Roux,
75724 Paris cedex 15, France

THE discovery of catalytically active RNA has provided the basis for the evolutionary concept of an RNA world. It has been proposed that during evolution the functions of ancient catalytic RNA were modulated by low molecular weight effectors, related to antibiotics, present in the primordial soup. Antibiotics and RNA may have coevolved in the formation of the modern ribosome¹. Here we report that a set of aminoglycoside antibiotics, which are known to interact with the decoding region of the 16S ribosomal RNA of *Escherichia coli*²⁻⁴, inhibit the second step of splicing of the T4 phage-derived *td* intron. Thus catalytic RNA seems to interact not only with a mononucleotide⁵ and an amino acid⁶, but also with another class of biomolecules, the sugars. Splicing of other group I introns but not group II introns was inhibited. The similarity in affinity and specificity of these antibiotics for group I introns and rRNAs may result from recognition of evolutionarily conserved structures.

The first step of splicing, which is initiated by binding of exogenous guanosine to the G-binding site, is dependent on guanosine concentration. Once bound to the ribozyme, the 5' splice-site is cleaved and the nucleoside becomes covalently linked to the first nucleotide of the intron. The second step of splicing is initiated by nucleophilic attack of the 3' splice site by the free 3' hydroxyl group of the upstream exon. The splice site is cleaved and the exons are ligated (Fig. 1)⁷. Unspliced

precursor RNA from the thymidylate synthase gene of the phage T4, which had been transcribed *in vitro*, and purified as previously described⁸, was incubated with increasing amounts of gentamicin under normal splicing conditions. Figure 2a (lanes 1-8) shows the inhibition of splicing at a GTP concentration of 2.5 μ M (the Michaelis constant (K_m) for GTP is 1 μ M⁸), whereas for lanes 9-16 the GTP concentration was 100 μ M. At concentrations ≤ 2 mM, gentamicin has little effect on the first step of splicing whatever the guanosine concentration. By contrast, the second step, cleavage at the 3' splice-site and exon ligation, is completely inhibited at a concentration of 2.5 μ M. The formation of ligated exons (E1-E2) and the appearance of a linear intron (L-In) are inhibited, whereas the intermediate products, intron-exon 2 (In-E2) and exon 1 (E1), remain nearly constant (Fig. 2a).

To analyse structure/activity relationships in the interaction of aminoglycosides with the ribozyme, we tested structurally related members of the families of aminoglycoside antibiotics (gentamicin, kanamycin and neomycin) in similar splicing inhibition assays (Fig. 2a). Tobramycin and neomycin B inhibited splicing at even lower concentrations than gentamicin, there being activity at a concentration as low as 0.5 μ M which also affected the first step of splicing, albeit only at an antibiotic concentration of 100 μ M. In Fig. 3a-c we list the antibiotics tested with their structures and the concentration at which about 50% inhibition was achieved. By comparing these results it becomes clear that the nature of the aminosugar is a factor in the inhibition of ribozyme function. To exclude nonspecific effects (for example, positive charge), different concentrations of spermidine were tested in combination with gentamicin, but spermidine had no influence on the inhibition caused by the antibiotic.

Other group I introns were tested for their sensitivity towards aminoglycosides. For the *Tetrahymena* rRNA, 50% inhibition of splicing occurs at ~ 100 μ M gentamicin. A shortened version of the *sunY* intron is sensitive to 10 μ M neomycin B and 100 μ M gentamicin. Figure 2b illustrates the sensitivity of the *sunY* intron towards gentamicin in a GTP-labelling experiment⁹. The lower sensitivity of the *sunY* and *Tetrahymena* introns towards both gentamicin and neomycin B could partially be explained by the stronger pairings of the P9.0 and P10 elements when compared with *td* (refs 10-12). As the first step of splicing in the *td* intron is also affected by the most potent antibiotic (neomycin B), we can expect the mechanism of inhibition for all aminoglycosides to be the same. As inhibition of splicing of the *sunY* intron by gentamicin and neomycin B takes place independently of the guanosine concentration, we believe that, in contrast to the inhibition of splicing by streptomycin^{13,14}, the antibiotic binds to a site distinct from the G-binding site¹⁰.

Inhibition in some cases (for example, tobramycin, neomycin B) occurred at concentrations equal to or lower than those required for inhibition of protein synthesis. Neomycin B is a

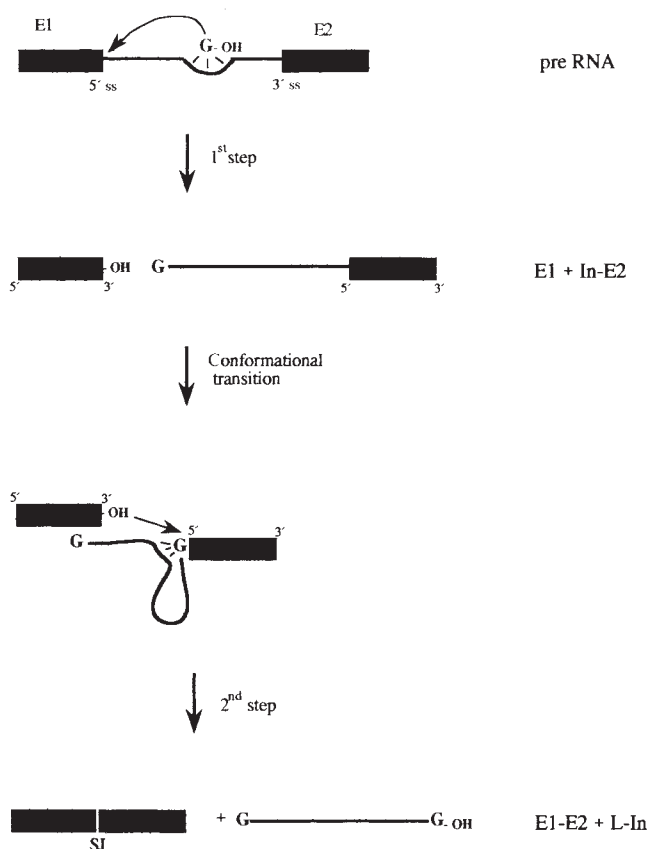


FIG. 1 Splicing pathway of group I introns. Solid bars indicate exons, lines indicate intron sequences; 5' ss, 3' ss and SJ indicate 5' and 3' splice sites and spliced junction, respectively. The first step of splicing is initiated by exogenous guanosine (G-OH) which attaches to the G-binding site, cleaves the 5' splice site by nucleophilic attack and becomes covalently added to the first nucleotide of the intron. For the second step of splicing the last nucleotide of the intron, always a G, has to enter the G-binding site (conformational change). The 3' OH group of the upstream exon attacks the 3' splice-site, which is cleaved and the exons ligated.

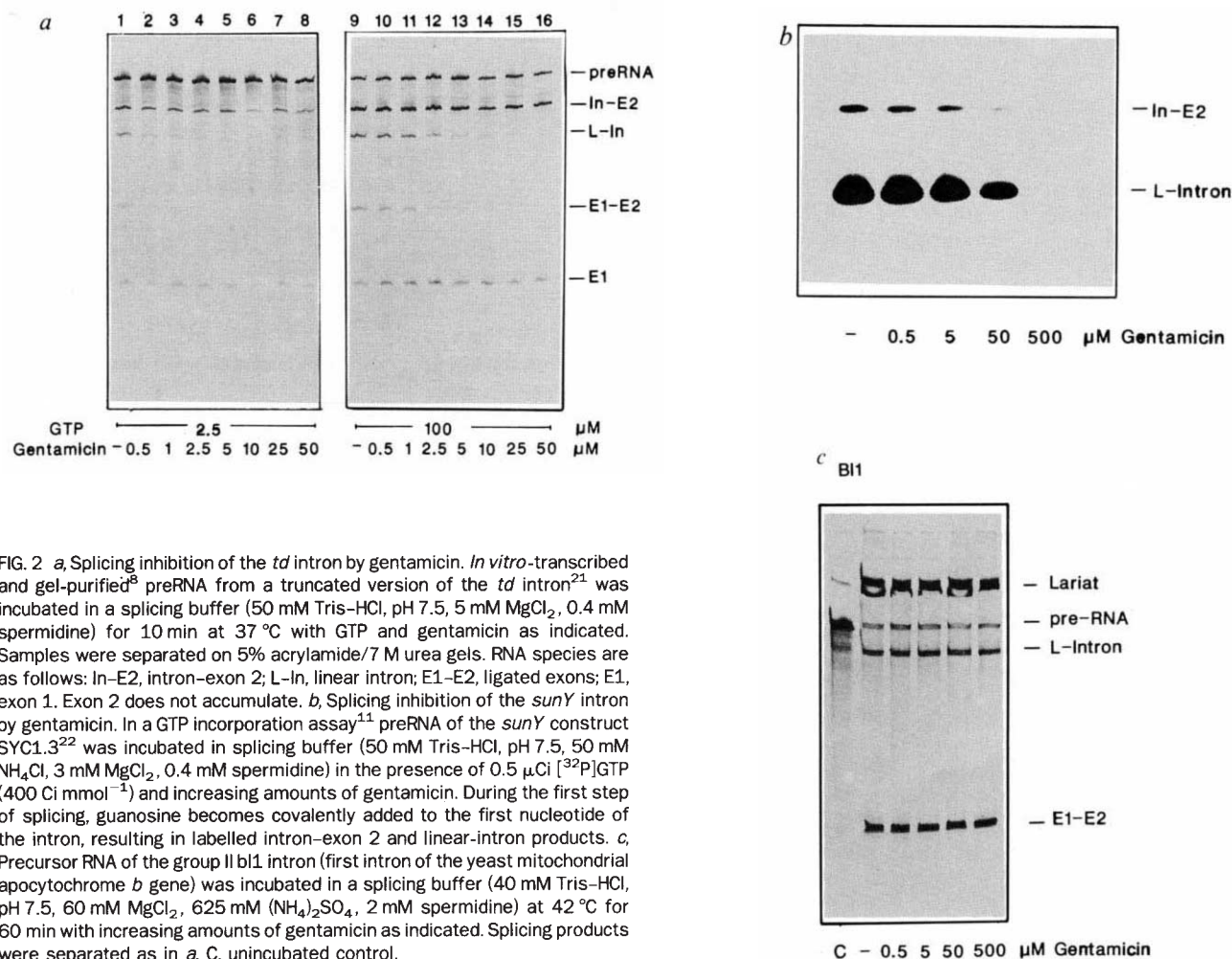


FIG. 2 *a*, Splicing inhibition of the *td* intron by gentamicin. *In vitro*-transcribed and gel-purified⁸ preRNA from a truncated version of the *td* intron²¹ was incubated in a splicing buffer (50 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 0.4 mM spermidine) for 10 min at 37 °C with GTP and gentamicin as indicated. Samples were separated on 5% acrylamide/7 M urea gels. RNA species are as follows: In-E2, intron-exon 2; L-In, linear intron; E1-E2, ligated exons; E1, exon 1. Exon 2 does not accumulate. *b*, Splicing inhibition of the *sunY* intron by gentamicin. In a GTP incorporation assay¹¹ preRNA of the *sunY* construct SYC1.3²² was incubated in splicing buffer (50 mM Tris-HCl, pH 7.5, 50 mM NH₄Cl, 3 mM MgCl₂, 0.4 mM spermidine) in the presence of 0.5 μCi [³²P]GTP (400 Ci mmol⁻¹) and increasing amounts of gentamicin. During the first step of splicing, guanosine becomes covalently added to the first nucleotide of the intron, resulting in labelled intron-exon 2 and linear-intron products. *c*, Precursor RNA of the group II b11 intron (first intron of the yeast mitochondrial apocytochrome *b* gene) was incubated in a splicing buffer (40 mM Tris-HCl, pH 7.5, 60 mM MgCl₂, 625 mM (NH₄)₂SO₄, 2 mM spermidine) at 42 °C for 60 min with increasing amounts of gentamicin as indicated. Splicing products were separated as in *a*. C, unincubated control.

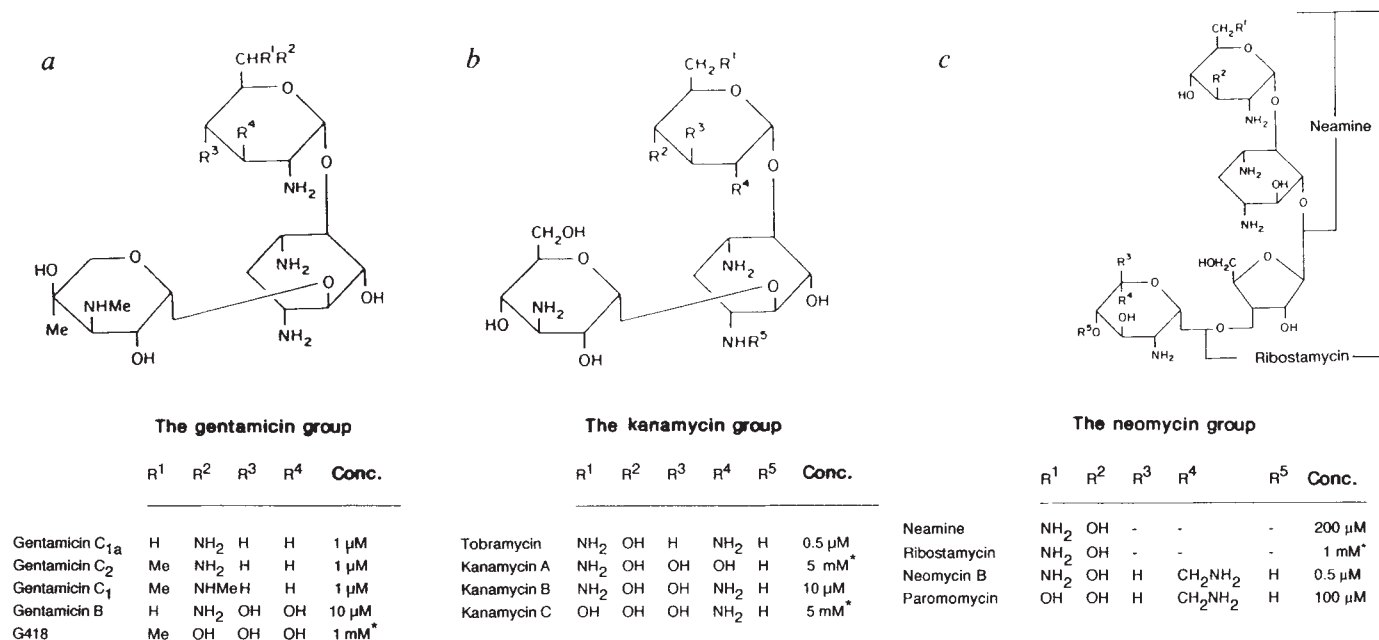


FIG. 3 *a-c*, Splicing inhibition experiments were done with the *td* intron as described in Fig. 2*a*, using several members of each family of aminoglycoside antibiotics (gentamicin, kanamycin and neomycin). Structures were taken

from ref. 23. Conc., antibiotic concentration at which about 50% inhibition was observed. * No inhibition was found up to the indicated concentration.

potent inhibitor of protein synthesis and the most active aminoglycoside in inhibiting the splicing reaction (Fig. 3). Comparison of the structurally related tobramycin, kanamycin A, B and C shows the order tobramycin > kanamycin B > kanamycin A > kanamycin C for splicing inhibition, and tobramycin equal to kanamycin B > kanamycin A > kanamycin C for inhibition of *in vitro* translation¹⁵. Positive charge does not seem to be a deciding factor and efficient inhibition of group I splicing is not a property of all antibiotics that inhibit translation: chloramphenicol, spectinomycin, hygromycin B, kasugamycin and erythromycin are inactive at millimolar concentrations (data not shown). The aminoglycosides do not generally inhibit catalytic RNA functions as splicing of a group II intron (the first intron of the yeast mitochondrial apocytochrome *b* gene) was not inhibited by gentamicin or by tobramycin at concentrations ≤ 5 mM (Fig. 2c).

The protein synthesis inhibitors shown to be active in group I splicing reactions have a common characteristic, namely induction of translation errors. They interact with the ribosome through specific interactions with sequences near the 3' end of 16S rRNA (as identified by mutation or protection studies), disturbing the fidelity of codon-anticodon recognition. The aminoglycoside-sensitive group I intron and the region of 16S rRNA implicated in antibiotic binding sites show no sequence homology; it is probable that RNA conformation, not sequence, determines antibiotic binding sites.

Several features of the group I introns suggest similarities in structure with the translational machinery. The guanosine-binding site of the group I introns is also an arginine-binding site and always an arginine codon^{5,6,10,16}. Furthermore, several proteins¹⁷⁻¹⁹ that are components of the translational apparatus promote group I intron splicing.

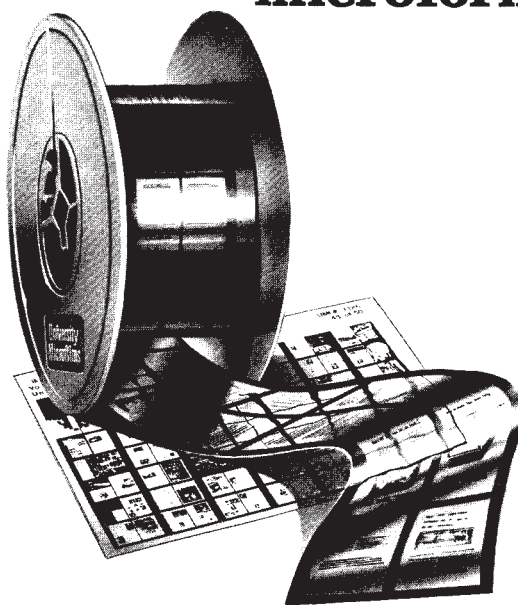
E. F. Gale has commented that new principles would emerge "from studies of the interactions of ribosomes or other forms of active RNA with streptomycin, chloramphenicol, tetracycline, etc." (ref. 20). To our knowledge this is the first report of antibiotic inhibition of a specific step in ribozyme function; other RNA-catalysed phosphodiester exchange reactions may be sensitive. It is not known if inhibition occurs *in vivo*, but the demonstration that ribozymes are targets for the aminoglycoside antibiotics may provide a means of detecting group I introns in microorganisms and also lead to the development of new classes of anti-fungal agents based on these structures. □

Received 12 June; accepted 29 July 1991.

1. Davies, J. *Molec. Microbiol.* **4**, 1227-1232 (1990).
2. Moazed, D. & Noller, H. F. *Nature* **327**, 389-394 (1987).
3. Noller, H. F. *A. Rev. Biochem.* **53**, 119-162 (1984).
4. Cundliffe, E. in *The Ribosome* (eds Hill, W. E. et al.) 479-490 (Am. Soc. Microbiol., Washington 1990).
5. Bass, B. & Cech, T. R. *Biochemistry* **25**, 4473-4477 (1986).
6. Yarus, M. *Science* **240**, 1751-1758 (1988).
7. Cech, T. R. *A. Rev. Biochem.* **59**, 543-568 (1990).
8. Schroeder, R., von Ahsen, U. & Belfort, M. *Biochemistry* **30**, 3295-3303 (1991).
9. Ehrenman, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5875-5879 (1986).
10. Michel, F., Hanna, M., Green, R., Bartel, D. P. & Szostak, J. W. *Nature* **342**, 391-395 (1989).
11. Burke, J. M., Eshetik, J. S., Burfeind, W. R. & King, J. *Nature* **344**, 80-82 (1990).
12. Michel, F. & Westhof, E. *J. molec. Biol.* **216**, 585-610 (1990).
13. Ahsen von, U. & Schroeder, R. *Nucleic Acids Res.* **19**, 2261-2265 (1991).
14. Ahsen von, U. & Schroeder, R. *Nature* **346**, 801 (1990).
15. Benveniste, R. & Davies, J. *Antimicrob. Agents Chemother.* **4**, 402-409 (1973).
16. Yarus, M. & Christian, E. L. *Nature* **342**, 349-350 (1990).
17. Akins, R. & Lambowitz, A. *Cell* **50**, 331-345 (1987).
18. Herbert, C., Labouesse, M., Dujardin, G. & Sionimski, P. P. *EMBO J.* **7**, 473-483 (1988).
19. Seraphin, B., Simon, M., Boulet, A. & Faye, G. *Nature* **337**, 84-87 (1989).
20. Gale, E. F. in *Biochemical Studies of Antimicrobial Drugs* (eds Newton, B. A. & Reynolds, P. E.) 1-21 (Cambridge University Press, Cambridge, 1966).
21. Salvo, J. L. G., Coetzee, T. & Belfort, M. *J. molec. Biol.* **211**, 537-549 (1990).
22. Michel, F., Netter, P., Xu, M.-Q. & Shub, D. *Genes Dev.* **4**, 777-788 (1990).
23. Davies, J. & Yagisawa, M. in *Biochemistry and Genetic Regulation of Commercially Important Antibiotics* (ed. Vining, L. C.) 329-354 (Addison-Wesley, London, 1983).

ACKNOWLEDGEMENTS. We thank M. Belfort, C. Thompson and F. Michel for discussion. This work was funded by the Austrian Fonds zur Förderung der Wissenschaft und Forschung.

nature is available in microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone () _____

University
Microfilms
International